Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

M. Bezem

Completeness of resolution by transfinite induction

Computer Science/Department of Software Technology Report CS-R8837 October

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Completeness of Resolution by Transfinite Induction

Marc Bezem
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

By a novel argument we unify completeness proofs of various strategies of (ground) resolution. The principle of proof we use is transfinite induction with respect to a well-founded ordering. The ordering reflects the intuition that resolvents are 'somehow' smaller than parent clauses. Moreover our exposition shows how redundancies in the completeness proofs are removed by choosing a better strategy. We use Zorn's Lemma in dealing with infinite sets of ground clauses. This permits a completeness proof of quantified resolution without using Herbrand's Theorem. Finally we give a proof theoretic argument, instead of the usual model theoretic argument, proving the completeness of resolution combined with subsumption and the deletion of tautologies.

1980 Mathematics Subject Classification: 03B05, 03B35, 68T15.
1987 CR Categories: I.2.3, F.4.1.
Key Words & Phrases: completeness, resolution, resolution strategies, transfinite induction, Herbrand's Theorem.
Note: The material for this article has been taken from the PRISMA document P346. The research is part of the PRISMA project (PaRallel Inference and Storage MAchine), a joint effort with Philips Research Eindhoven, partially supported by the Dutch "Stimulerings-projectteam Informatica-onderzoek" (SPIN).

1. Preliminaries

1.1. Logic

Let $\mathfrak{L}=\{p_0,p_1,...,\}$ be an infinite set of propositional atoms. A *literal* is an atom or the negation of an atom. The literals p_i and $\neg p_i$ are called *complementary* (p_i is called *positive*, $\neg p_i$ *negative*). If L is a literal, than its complement is denoted by $\bar{L}$. The set of literals will be denoted by $\mathfrak{Lit}$. An *interpretation* is a subset I of $\mathfrak{L}$, corresponding to the truth valuation $\mathcal{V}_I(p_i)=$ TRUE if $p_i\in I$, and FALSE otherwise. A *clause* is a finite set of literals.

Truth of a literal L (respectively a clause C) in an interpretation I, denoted by $I\models L$ (respectively $I\models C$), is defined as follows:

$$I\models p_i \text{ iff } p_i\in I$$

$$I\models\neg p_i \text{ iff } p_i\notin I$$

$$I\models\{L_1,...,L_n\}\ (n\geqslant 0) \text{ iff } I\models L_i \text{ for some } 1\leqslant i\leqslant n$$

Note that the empty clause is false in any interpretation. The *truth set* (respectively *falsity set*) of an interpretation I is defined as $\mathfrak{T}(I)=\{L\in\mathfrak{Lit}\mid I\models L\}$ (respectively $\mathfrak{F}(I)=\{L\in\mathfrak{Lit}\mid$ not $I\models L\}=\mathfrak{Lit}-\mathfrak{T}(I)$). Note that neither $\mathfrak{T}(I)$ nor $\mathfrak{F}(I)$ contains complementary literals, and that both contain an occurrence, positive or negative, of every $p_i\in\mathfrak{L}$. For a set of clauses S we define $S_{\mathfrak{T}}(I)=\{C\in S\mid I\models C\}$ and $S_{\mathfrak{F}}(I)=S-S_{\mathfrak{T}}(I)$. A set of clauses S is called *consistent* if there exists an interpretation I in which every clause from S is true. Such an interpretation I is called a *model* of S (or: a model *for* S).

Report CS-R8837
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

1.2. Transfinite induction

Let $<$ be a well-founded partial ordering on a set S (i.e. $<$ is an irreflexive, transitive relation on S such that every descending sequence is finite). The principle of *transfinite induction* with respect to $<$ states that if some property P of elements of S is progressive, then it holds for every element of S. Here *progressivity* of P means that $P(x)$ is implied by $\forall y{<}x\ P(y)$, for all $x \in S$. The validity of this principle of proof, formally requiring an application of the axiom of choice, is intuitively obvious. For, assume P is progressive and $\neg P(x_0)$ for some $x_0 \in S$ (towards a contradiction). Then there exists $x_1 < x_0$ with $\neg P(x_1)$. Iteration of this argument yields an infinite descending sequence, which contradicts the well-foundedness of $<$. It should be noted that we do not use at all the full proof theoretic strength of transfinite induction. This strength can be measured by assigning ordinals to elements of S in the usual way: $\|x\| = sup\{\|y\|+1 \mid y<x\}$ (with $\sup \varnothing = 0$). We only use transfinite induction with respect to orderings having the property that $\|x\| < \omega$ (the order type of the natural numbers) for all $x \in S$. Whether one prefers transfinite induction or induction on the natural number $\|x\|$ (or even an informal argument in which the induction is not explicit) appears to be a matter of taste. However, for the purpose of unifying completeness proofs of various strategies of resolution and analyzing their differences, transfinite induction suits best.

1.3. Zorn's Lemma

Let $\subset$ be a partial ordering on a set S. A *chain* in S is a totally ordered subset of S (i.e. satisfying the trichotomy axiom). *Zorn's Lemma* (see [H]) states that S contains a maximal (minimal) element, provided that every chain in S has an upper (lower) bound in S. Zorn's Lemma is known to be one of the most practical equivalents of the axiom of choice. We use it in dealing with infinite sets of clauses.

1.4. Resolution

Resolution is the rule according to which a *resolvent* $R = (C - \{L\}) \cup (C' - \{\bar{L}\})$ may be inferred from *parent clauses* C and C', containing literals L and $\bar{L}$ respectively, and satisfying the requirements of the strategy. A *strategy* is, intuitively speaking, a prescription telling which clauses may be resolved. For example we can require that one of the clauses consists entirely of positive literals. For some strategies this prescription may extend over more than one inference. For our exposition this informal notion of strategy suffices.

A *derivation* (relative to a strategy) of a clause C from a set of clauses S is a sequence of clauses $C_1,...,C_n$ such that $C_n = C$ and, for all $1 \leqslant k \leqslant n$, either C_k is in S or C_k is a resolvent of some C_i and C_j with $1 \leqslant i,j < k$, provided that C_k may be inferred from C_i and C_j according to the strategy. For some strategies the notion of derivation has to be generalized by allowing C_k to be inferred from $C_{i_1},...,C_{i_j}$ with $1 \leqslant i_1,...,i_j < k$, $j \geqslant 2$, instead of just from C_i and C_j.

Resolution can easily be proved *sound*, i.e. for any interpretation I, resolution preserves truth in I. For, at least one of two true parent clauses must contain a true literal different from the two complementary literals which are resolved out. In particular consistency is preserved when resolvents are added to a set of clauses; inconsistency is preserved by definition.

Completeness of a resolution strategy is the property that from any inconsistent set of clauses the empty clause can be derived. Completeness is usually proved as follows: let S be an inconsistent set of clauses, close S under resolution according to the stategy and then apply the following

PROPOSITION. *Every set of clauses which is closed under resolution according to the stategy and does not contain the empty clause is consistent.*

In the next section we shall prove this proposition for various resolution strategies: binary resolution [R1], semantic resolution [M] (where the notion of renaming is introduced), P_1-resolution and hyperresolution [R2], ordered hyperresolution [S] (where the idea of ordering the atoms is attributed to Reynolds and worked out in the more general setting of semantic resolution), as well as SLD-resolution [K]. All proofs will be by transfinite induction and have the following general form: first prove for some well-founded ordering $<$ on S and some interpretation J that truth in J of clauses from S is progressive, then conclude that J is a model for S by transfinite induction.

2. Completeness

2.1. Binary resolution

In the case of *binary* resolution, no restrictions are specified and every two clauses containing complementary literals may be resolved. Let S be a set of clauses which is closed under binary resolution and does not contain the empty clause. Note that we do not assume that S is finite. Fix an arbitrary interpretation I. The interpretation I may informally be seen as a first try for a model of S. If this try fails (i.e. if $S_{\mathfrak{F}}(I)$ is not empty), then I has to be 'adjusted' on literals occurring in clauses from $S_{\mathfrak{F}}(I)$. This adjustment should not affect the truth of other clauses. Therefore some 'minimal' adjustment is made, yielding a model J of S (this idea goes back to [R2]).

Let X be the set of all literals which occur in clauses from $S_{\mathfrak{F}}(I)$ (formally $X = \cup S_{\mathfrak{F}}(I)$). Note that $X \subseteq \mathfrak{F}(I)$ does not contain complementary literals. We say that a subset Y of X *covers* $S_{\mathfrak{F}}(I)$ (or: Y is a *covering* of $S_{\mathfrak{F}}(I)$) if every clause from $S_{\mathfrak{F}}(I)$ contains at least one literal from Y. For example X itself covers $S_{\mathfrak{F}}(I)$, since S does not contain the empty clause. We shall construct a minimal (wrt. set inclusion) subset Y of X that covers $S_{\mathfrak{F}}(I)$. If X is finite, then Y is easily obtained from X by deleting elements in such a way that the resulting set still covers $S_{\mathfrak{F}}(I)$. If X is infinite, then this process is iterated, intuitively speaking, in a transfinite way until eventually Y is reached:

$$X_0 = X, \ldots, X_{i+1} = X_i - \{x\}, \ldots, X_\omega = \bigcap_{i<\omega} X_i,\ X_{\omega+1} = X_\omega - \{x'\}, \ldots,$$

with $x \in X_i$ such that X_{i+1} covers $S_{\mathfrak{F}}(I)$, and $x' \in X_\omega, \ldots$ similarly. Note that we tacitly assumed that, for example, X_ω covers $S_{\mathfrak{F}}(I)$. As there are much more ordinals in the universe than elements of X, this process terminates with a minimal set covering $S_{\mathfrak{F}}(I)$.

This informal argument can be made rigorous by applying Zorn's Lemma. Let Z be the set of subsets of X that cover $S_{\mathfrak{F}}(I)$. Z is partially ordered by set inclusion. Existence of a minimal Y in Z is guaranteed by Zorn's Lemma if we prove that every chain in Z has lower bound in Z. Let Z' be a chain in Z. The set $\cap Z'$ (with $\cap \varnothing = X$) is certainly a lower bound of Z', so it suffices to prove that $\cap Z'$ is in Z, i.e. covers $S_{\mathfrak{F}}(I)$. Suppose $\{L_1,\ldots,L_n\}$ is a clause in $S_{\mathfrak{F}}(I)$ having no literal in common with $\cap Z'$ (towards a contradiction). Then there exists for every $1 \leqslant i \leqslant n$ an element, say X_i, of Z' which does not contain L_i. Since Z' is a chain, the X_i's are totally ordered. Hence some X_i is a subset of all of them, and hence contains none of the literals $L_1,\ldots,L_n$. This clearly contradicts $X_i \in Z' \subseteq Z$ by the definition of Z.

Given a minimal set Y covering $S_{\mathfrak{F}}(I)$ we define J to be the (unique) interpretation such that $\mathfrak{F}(I) \cap \mathfrak{T}(J) = Y$ (formally $J = \{p_i \in \mathfrak{L} \mid p_i \in Y \vee (\neg p_i \notin Y \wedge p_i \in I)\}$). In other words: the interpretation J is such that the truth valuations $\mathfrak{V}_J$ and $\mathfrak{V}_I$ only differ on the atoms which occur, positively or negatively, in Y. Since Y is a minimal covering of $S_{\mathfrak{F}}(I)$ it follows that J is a model of $S_{\mathfrak{F}}(I)$ having the property that for every literal $L \in \mathfrak{T}(J)$ which occurs in a clause from $S_{\mathfrak{F}}(I)$ there exists a clause in $S_{\mathfrak{F}}(I)$ in which L is the *only* literal from $\mathfrak{T}(J)$. This property of J is crucial and shall be used in the proof of the lemma below.

We now arrive at the point where the ordering $<$ is defined. Let $<$ be the transitive closure of the relation $<_1$ on S defined by

$$R <_1 C \text{ iff } R \text{ is the resolvent of } C \text{ and some } C' \in S_{\mathfrak{F}}(I).$$

As R contains less literals from $\mathfrak{F}(I)$ than C (recall that $\cup S_{\mathfrak{F}}(I) \subseteq \mathfrak{F}(I)$), it follows that $<$ is a well-founded partial ordering. The lemma below implies that truth in J is progressive. It follows by transfinite induction that J is a model of S, and hence S is consistent. This completes the proof of the proposition in 1.4 in the case of binary resolution.

LEMMA. *For every C in S we have: if $\forall R <_1 C\ J \models R$, then $J \models C$*

PROOF. Let C be a clause of S such that $\forall R <_1 C\ J \models R$. If $C \in S_{\mathcal{F}}(I)$, then we immediately have $J \models C$ since J is a model of $S_{\mathcal{F}}(I)$. Now assume $C \in S_{\mathcal{T}}(I)$ is false in J (towards a contradiction), then C consists entirely of literals from $\overline{\mathcal{T}(J)}$. Since C is true in I, it follows that C contains a literal $L \in \mathcal{T}(I) \cap \overline{\mathcal{T}(J)}$, so $\overline{L} \in \overline{\mathcal{T}(I)} \cap \mathcal{T}(J) = Y$. Now by the crucial property of J stated above there exists a clause $C' \in S_{\mathcal{F}}(I)$ such that $\overline{L}$ is the only literal of C' which is true in J. Hence $R = (C - \{L\}) \cap (C' - \{\overline{L}\}) <_1 C$ and R consists entirely of literals which are false in J. This clearly contradicts $\forall R <_1 C\ J \models R$. □

2.2. Digression: elimination of Herbrand's Theorem

In the previous subsection we obtained, by using Zorn's Lemma, a completeness result for arbitrary, not necessarily finite, sets of (ground) clauses. We can exploit this completeness result by proving the completeness of quantified resolution without appealing to Herbrand's Theorem. To this end we adopt, for the time of this subsection, the terminology and notations of [R1]. We prepare by the following two lemmas, where *ground*(S) denotes the set of ground instances of clauses from S.

LEMMA 2.2.1. *A set of clauses S is consistent if ground(S) is consistent.*

PROOF. A model of *ground*(S) is a Herbrand model of S. □

LEMMA 2.2.2. *The empty clause is derivable from a set of clauses S if it is derivable from ground(S).*

PROOF. By the Lifting Lemma from [R1, 5.15]. □

Now we can easily prove the desired completeness result. Let S be an inconsistent set of clauses. Then *ground*(S) is inconsistent by Lemma 2.2.1. From the completeness of ground resolution it follows that the closure of *ground*(S) under binary resolution contains the empty clause, i.e. the empty clause is derivable from *ground*(S). Hence by Lemma 2.2.2 the empty clause is derivable from S. This completes the proof of the completeness of quantified resolution.

2.3. Redundancy

If one takes closer look at the argument developed in 2.1, then the following observations can be made:

- The interpretation I on which the argument is based is arbitrary;
- The minimal set Y covering $S_{\mathcal{F}}(I)$ may not be unique;
- Lemma 2.1 is stronger than progressivity since $\forall R <_1 C\ J \models R$ is weaker than $\forall R < C\ J \models R$.

These observations reveal substantial redundancies in the completeness proof, since for any interpretation I, any minimal Y covering $S_{\mathcal{F}}(I)$, and even with $<_1 = <$ a completeness result can be obtained.

In general, a resolution strategy aims at reducing the costs of finding a derivation of the empty clause from a given set of clauses S. If a strategy is complete, then we can simply close S under resolution according to the strategy, until eventually the empty clause is derived. The costs of this closing procedure are determined by the number of generated resolvents. Thus the importance of reducing the number of generated resolvents becomes evident. To this end various strategies of resolution exploit the redundancies in the completeness proof of 2.1 mentioned above: semantic resolution (with P_1-resolution and SLD-resolution as special cases) fixes I, hyperresolution fixes I and trivializes the ordering ('$<_1 = <$'), whereas ordered hyperresolution exploits the non-uniqueness of Y as well. We shall discuss these matters in the following subsections.

2.4. Semantic resolution

In the case of *semantic* resolution, an interpretation I is fixed in advance. Given a set of clauses S, resolution is only allowed between a clause from $S_{\mathfrak{T}}(I)$ and one from $S_{\mathfrak{F}}(I)$. This restriction does not at all affect the completeness proof from 2.1. Hence semantic resolution is complete.

P_1-resolution [R2] is obtained as a special case of semantic resolution by taking $I = \varnothing$. Then $S_{\mathfrak{F}}(I)$ consists of the clauses from S not containing negated atoms, so-called *positive* clauses.

SLD-resolution [K] is a rule of inference for so-called Horn clauses. A *Horn clause* is a clause with at most one positive literal. Note that the set of all Horn clauses is closed under binary resolution. We distinguish between *program clauses* (or *definite clauses*), which contain exactly one positive literal, and *goal clauses*, which consist entirely of negated atoms. Thus the empty clause is a goal clause. SLD-resolution uses a *selection rule*, which selects from every goal clause a (negative) literal. Resolution is only allowed between program clauses and goal clauses, and with the restriction that the negation of the positive literal of the program clause is the selected literal of the goal clause. SLD-resolution can be viewed as semantic resolution with $I = \mathfrak{L}$: for a set of Horn clauses S, $S_{\mathfrak{F}}(\mathfrak{L})$ consists of the goal clauses from S, and $S_{\mathfrak{T}}(\mathfrak{L})$ of the program clauses. With some technical effort (concerning selection rules) the completeness of SLD-resolution can be obtained from the completeness of semantic resolution. We refrain from giving a detailed account on this point.

As done in [S], hyperresolution as well as ordered hyperresolution (and also SLD-resolution) could be treated more generally in the context of semantic resolution. For reasons of simplicity, however, we prefer to specialize to the case $I = \varnothing$. Modulo renaming from [M] we do not loose generality.

2.5. Hyperresolution

In [R2] hyperresolution was introduced as a refinement of P_1-resolution. A hyperresolvent of a set of clauses S is a positive clause which is obtained by successive P_1-resolutions in a way depicted in Figure 1. More precisely: a positive clause C_{n+1} is called a *hyperresolvent* of S with parent clause C_1 if $n \geqslant 1$, $C_1 \in S$, $D_i \in S$ is positive and C_{i+1} is a P_1-resolvent of C_i and D_i, for all $1 \leqslant i \leqslant n$.

$$
\begin{array}{ll}
C_1 & D_1 \\
\mid & / \\
C_2 & D_2 \\
\mid & / \\
\vdots & \vdots \\
C_n & D_n \\
\mid & / \\
C_{n+1} &
\end{array}
$$

FIGURE 1.

If we assume that S is closed under hyperresolution, then we can define a relation $<_1^h$ on S by

$$R <_1^h C \quad \text{iff } R \text{ is hyperresolvent of } S \text{ with parent clause } C.$$

Since every hyperresolvent is positive and can as such not act as a parent clause of another hyperresolvent, it follows trivially that $<_1^h$ is a well-founded partial ordering. As $<_1^h$ equals its transitive closure $<^h$ we shall drop the subscript.

In order to compare $<^h$ with $<$ we assume for the time of this paragraph that S is closed under P_1-resolution, which implies closure under hyperresolution. We then have

$$R <^h C \quad \text{iff } R < C \text{ and } R \text{ is positive.}$$

Since positive clauses are $<$-minimal we can view $<^h$ as 'cutting short' $<$.

Let S be a set of clauses which is closed under hyperresolution and does not contain the empty clause. The argument that S is consistent is again similar to that given in 2.1, taking $I=\varnothing$, $<^h_{(1)}$ instead of $<_{(1)}$, reading 'hyperresolution' for 'resolution', and so on. Only the proof of the progressivity of truth in the interpretation J needs some more attention.

LEMMA. *Under conditions as above we have for every C in S: if $\forall R <^h C\ J \models R$, then $J \models C$.*

PROOF. Let C be a clause of S such that $\forall R <^h C\ J \models R$. If $C \in S_{\mathfrak{T}}(\varnothing)$ (i.e. C is positive), then we immediately have $J \models C$ since J is a model of $S_{\mathfrak{T}}(\varnothing)$. Now assume $C \in S_{\mathfrak{T}}(\varnothing)$ is false in J, i.e. C consists entirely of literals which are false in J (towards a contradiction). Let $L_1,...,L_n (n>0)$ be the negative literals of C (here we deviate from 2.1 if $n>1$). We have $L_1,...,L_n \in \mathfrak{T}(\varnothing) \cap \mathfrak{T}(J)$, so $\overline{L}_1,...,\overline{L}_n \in \mathfrak{T}(\varnothing) \cap \mathfrak{T}(J) = Y$. It follows by the minimality of Y that there exist $D_1,...,D_n \in S_{\mathfrak{T}}(\varnothing)$ (i.e. positive clauses) such that $\overline{L}_i$ is the *only* literal of D_i which is true in J $(1 \leqslant i \leqslant n)$. Hence the hyperresolvent R of S with parent clause C, obtained as in Figure 2, consists entirely of literals which are false in J. This contradicts $\forall R <^h C\ J \models R$. □

$$
\begin{array}{rcc}
C_1 = & C & D_1 \\
 & | & / \\
C_2 = & (C_1 - \{L_1\}) \cup (D_1 - \{\overline{L}_1\}) & D_2 \\
 & | & / \\
C_3 = & (C_2 - \{L_2\}) \cup (D_2 - \{\overline{L}_2\} & D_3 \\
 & | & / \\
 & \vdots & \vdots \\
 & C_n & D_n \\
 & | & / \\
 & R &
\end{array}
$$

FIGURE 2.

2.6. Compactness

The remaining completeness proof (for ordered hyperresolution) shall be given for finite sets of clauses only. However, on the basis of our previous results, the finite case can easily be generalized to the infinite case. This will be achieved by the following version of the Compactness Theorem.

THEOREM. *Every inconsistent set of clauses has a finite subset which is inconsistent.*

PROOF. Let S be an inconsistent set of clauses. By the definition of consistency the closure T of S under binary resolution is also inconsistent. By the completeness of binary resolution the set T must contain the empty clause. Hence there exists a derivation of the empty clause from S by binary resolution. As derivations are finite, it follows that there exists a derivation of the empty clause from a finite subset of S. Now by the soundness of resolution this finite subset of S is inconsistent. □

COROLLARY. *Assume a strategy which is monotonic in the sense that every derivation from a set of clauses is also a derivation from any superset of that set of clauses. Assume furthermore that every finite set of clauses, which is closed under resolution according to the strategy and does not contain the empty clause, is consistent. Then this latter property also holds for infinite sets of clauses.*

The remaining completeness proof has the following form: let S be an inconsistent finite set of clauses, then the closure of S under resolution according to the stategy is also finite and inconsistent, and hence contains the empty clause by the following proposition.

PROPOSITION. *Every finite set of clauses which is closed under resolution according to the stategy and does not contain the empty clause is consistent.*

2.7. Ordered hyperresolution

In the case of ordered hyperresolution (see [S]), the atoms in $\mathcal{L} = \{p_0, p_1, ..., p_m\}$ are totally ordered. Recall Figure 1 and assume $C_{i+1} = (C_i - \{\neg p_{k_i}\}) \cup (D_i - \{p_{k_i}\})$ for $1 \leqslant i \leqslant n$. For C_{n+1} to be an *ordered* hyperresolvent we require p_{k_i} to be the *maximal* atom of D_i, for all $1 \leqslant i \leqslant n$. It will be clear that ordered hyperresolution is a monotonic strategy in the sense of Corollary 2.6.

Ordered hyperresolution is seen to be complete in the same way as hyperresolution. The restriction upon the p_{k_i}'s has the effect that the argument in 2.1 should be modified as follows. Instead of starting with a covering set $X = \cup S_{\mathcal{G}}(\emptyset)$ we start with $X_0 = \{p_i \mid p_i$ is the maximal atom of a clause in $S_{\mathcal{G}}(\emptyset)\}$. Furthermore, the deletion process should be such that the minimal atom $p_j \in X_i$ such that $X_i - \{p_j\}$ covers $S_{\mathcal{G}}(\emptyset)$ is deleted. So we have formally $X_{i+1} = X_i - \{min\{p_j \mid p_j \in X_i$ and $X_i - \{p_j\}$ covers $S_{\mathcal{G}}(\emptyset)\}\}$. Here we use finiteness, since otherwise such a minimal atom does not always exist. Thus in a finite number of steps a minimal set Y covering $S_{\mathcal{G}}(\emptyset)$ is obtained, having the property that for every $p_i \in Y$ there exists a clause in $S_{\mathcal{G}}(\emptyset)$ in which p_i is the *maximal* and the *only* literal from Y. For, there exist clauses in $S_{\mathcal{G}}(\emptyset)$ in which p_i is the only literal from Y (otherwise Y would not be a minimal covering). Assume all such clauses have a maximum greater than p_i (towards a contradiction). Before deletion of one of these maxima, deletion of p_i would also result in a set which still covers $S_{\mathcal{G}}(\emptyset)$. But then the deletion of the first of these maxima would be contrary to the definition of the deletion process, since p_i is smaller than the deleted literal. By this contradiction we have proved the desired property of the minimal covering Y.

2.8. Resolution with subsumption and the deletion of tautologies

In the preceding subsections we unified completeness proofs of various strategies of resolution. We do not claim to be exhaustive, not even with respect to strategies that fit into our unifying framework. Of all possible refinements we just mention two obvious improvements, both concerning the deletion of redundant clauses. We shall restrict ourselves to binary resolution, but the results can be generalized to any other strategy that we dealt with.

A clause which contains complementary literals is called a *tautology*. A tautology is true in every interpretation. For clauses C, C' with $C \subseteq C'$ we say that C *subsumes* C'. Obviously C' is true in every interpretation in which C is true. The notion of subsumption first appeared in [R1]. For any set of clauses S, let $nontaut(S)$ denote the set of clauses from S which are not tautologies, $min(S)$ the set of clauses in S which are not subsumed by any other clause from S (these clauses are minimal in S with respect to set inclusion), and $res(S)$ the set of clauses containing S as well as all resolvents of clauses from S. We view *nontaut*, *min* and *res* as operators on sets of clauses; composition will be denoted by $\circ$ and iteration by an exponent. It is easy to see that *nontaut* and *min* are idempotent and commute (since tautologies only subsume tautologies).

Completeness of resolution with subsumption is usually obtained with the help of the following model theoretic theorem, whose obvious proof we omit.

SUBSUMPTION THEOREM. *A set of clauses S is consistent if and only if min(S) is consistent.*

Completeness is then proved by showing that every set of clauses S, which does not contain the empty clause and has the property that every resolvent of clauses from S is subsumed by an element of S, is consistent. As long as the sets of clauses are finite this argument is sufficient. However, in the case of infinite sets of clauses the fact that subsumption is a non-monotonic operation complicates the closure procedure. To avoid these difficulties we present proof theoretic arguments for the completeness of resolution combined with subsumption and the deletion of tautologies.

We need a notion called *rank* of a clause, expressing intuitively the least depth of a proof tree for that clause. Let $C_1,...,C_n$ be a derivation from a set of clauses S, and let T be the set of clauses occurring in this derivation. We define disjoint subsets T_i of T in the following inductive way: $T_0=T\cap S$; $T_{i+1}=(res(T_0\cup\cdots\cup T_i)\cap T)-(T_0\cup\cdots\cup T_i)$. Since T is finite there exists a least number k for which $T_{k+1}=\varnothing$; T equals the disjoint union of $T_0,...,T_k$ for this k. Now the *rank* of C_i in the derivation $C_1,...,C_n$ is the unique number j with $C_i\in T_j$. (Note that the rank of a clause in a derivation may be based on a reordering of that derivation.) The *rank* of a derivation is the maximum rank of its clauses. We call a derivation *normalized* if the ranks of its clauses are non-decreasing. Every derivation can be reordered into a normalized derivation. If a derivation from S has rank k, then its clauses are in $res^k(S)$. Conversely, every $C\in res^k(S)$ has a derivation from S of rank at most k.

SUBSUMPTION LEMMA. *For every derivation $C_1,...,C_n$ from a set of clauses S there exists a derivation $C'_1,...,C'_n$ from $min(S)$ such that C'_i subsumes C_i for every $1\leqslant i\leqslant n$. Moreover, the rank of $C'_1,...,C'_n$ is at most k if $C_1,...,C_n$ is a normalized derivation of rank k.*

PROOF. By induction on the length of the derivation. In the base case $n=1$ we have $C_1\in S$ and can take $C'_1\in min(S)$ subsuming C_1. As to the induction step, assume the Subsumption Lemma has been proved for derivations of length n and let $C_1,...,C_{n+1}$ be a derivation from S. By the induction hypothesis there exist $C'_1,...,C'_n$ having the desired properties. If $C_{n+1}\in S$ then we proceed as in the base case. If C_{n+1} is a resolvent of some C_i and C_j with $1\leqslant i,j\leqslant n$, then $C_{n+1}=(C_i-\{L\})\cup(C_j-\{\bar{L}\})$ for some literal L with $L\in C_i$ and $\bar{L}\in C_j$. By the induction hypothesis we have $C'_i\subseteq C_i$ and $C'_j\subseteq C_j$. We distinguish the following four cases. (*i*) $L\in C'_i$ and $\bar{L}\in C'_j$. Then take $C'_{n+1}=(C'_i-\{L\})\cup(C'_j-\{\bar{L}\})$. (*ii*) $L\in C'_i$ and $\bar{L}\notin C'_j$. Then $C'_j\subseteq C_{n+1}$, so take $C'_{n+1}=C'_j$. (*iii*) $L\notin C'_i$ and $\bar{L}\in C'_j$. Then $C'_i\subseteq C_{n+1}$, so take $C'_{n+1}=C'_i$. (*iv*) $L\notin C'_i$ and $\bar{L}\notin C'_j$. Then $C'_i,C'_j\subseteq C_{n+1}$, so take (arbitrarily) $C'_{n+1}=C'_i$. In all cases we have $C'_{n+1}\subseteq C_{n+1}$. It should be obvious that $C'_1,...,C'_{n+1}$ is a derivation from $min(S)$ having the desired properties. □

TAUTOLOGY LEMMA. *For every derivation $C_1,...,C_n$ from a set of clauses S there exists a sequence of clauses $C'_1,...,C'_n$ such that C'_i subsumes C_i and $C'_i=C_i$ if C_i is a tautology ($1\leqslant i\leqslant n$), and the sequence obtained by deleting every tautology from $C'_1,...,C'_n$ is a derivation from $nontaut(S)$. Moreover this last derivation has rank at most k if $C_1,...,C_n$ is a normalized derivation of rank k.*

PROOF. By induction on the length of the derivation. In the base case $n=1$ we can simply take $C'_1=C_1$. As to the induction step, assume the Tautology Lemma has been proved for derivations of length n and let $C_1,...,C_{n+1}$ be a derivation from S. By the induction hypothesis there exist $C'_1,...,C'_{n+1}$ having the desired properties. If $C_{n+1}\in S$ or C_{n+1} is a tautology, then we proceed as in the base case. If neither $C_{n+1}\in S$ nor C_{n+1} is a tautology, then C_{n+1} is a resolvent of some C_i and C_j with $1\leqslant i,j\leqslant n$ which are not both tautologies (a resolvent of two tautologies must itself be a tautology). We distinguish the following two cases. (1) Neither C_i nor C_j is a tautology. Then this also holds for C'_i and C'_j and we proceed as in the cases (i)-(iv) of the proof of the Subsumption Lemma. (2) One of C_i and C_j is a tautology. Then by symmetry we may assume that C_i is a tautology and C_j is not. Since C_{n+1} is not a tautology we must have that $C_{n+1}=(C_i-\{L\})\cup(C_j-\{\bar{L}\})$ for some literal L such that $L,\bar{L}\in C_i$ and $\bar{L}\in C_j$. It follows that $C_j\subseteq C_{n+1}$. Now take $C'_{n+1}=C'_j$. In all cases we have $C'_{n+1}\subseteq C_{n+1}$. After a moment's reflection one sees that the sequence $C'_1,...,C'_{n+1}$ has the desired properties. □

The above lemmas allow us to infer the completeness of resolution with subsumption and the deletion of tautologies from the completeness of resolution. More precisely, we can find a derivation of the empty clause from any inconsistent set of clauses by applying at random, but fair with respect to *res*, the operations *nontaut*, *min* and *res*. We shall illustrate this with a typical example. Let S be an inconsistent set of clauses. Then by the completeness of resolution there exists a least number k, say $k=3$, such that the empty clause occurs in $res^k(S)$. We show that the empty clause also occurs in $(res \circ min \circ nontaut \circ res^2 \circ min)(S)$. There exists a normalized derivation of the empty clause from S of rank 3. So by the Subsumption Lemma there exists a derivation of the empty clause from $min(S)$ of rank at most 3. Hence the empty clause occurs in $res^3(min(S)) = res((res^2 \circ min)(S))$. Now it follows by subsequent application of the Tautology Lemma and the Subsumption Lemma that the empty clause occurs in $(res \circ min \circ nontaut \circ res^2 \circ min)(S)$.

ACKNOWLEDGEMENT.
This paper has benefitted much from discussions with Anton Eliëns and Jan Willem Klop.

REFERENCES

[H] P.R. HALMOS, (1960). *Naive set theory.* D. van Nostrand Company, Princeton, p. 62.
[K] R.A. KOWALSKI, (1974). *Predicate logic as a programming language.* In: J. ROSENFELD (editor), *Information Processing 74,* Stockholm, North-Holland, Amsterdam, pp. 569-574.
[M] B. MELTZER, (1966). *Theorem-proving for computers: some results on resolution and renaming.* Computer Journal 8, pp. 341-343.
[R1] J.A. ROBINSON, (1965). *A machine oriented logic based on the resolution principle.* Journal of the ACM 12, 1, pp. 23-41.
[R2] J.A. ROBINSON, (1965). *Automatic deduction with hyper-resolution.* International Journal of Computer Mathematics 1, pp. 227-234.
[S] J.R. SLAGLE, (1967). *Automatic theorem proving with renamable and semantic resolution.* Journal of the ACM 14, 4, pp. 687-697.