Tautological axioms in tree-like k-DNF Resolution

Jan Johannsen

February 25, 2022

k-DNF Resolution is a generalization of Resolution introduced by Krajíček [2] in the context of Bounded Arithmetic, which uses k-DNF formulas instead of just clauses. Cut-elimination shows that the tree-like versions $Res^*(k)$ of these systems are actually weaker than dag-like Resolution [1], so that they form a hierarchy of proof systems betweeen tree-like and dag-like resolution.

It is well-known that tautological axioms are redundant for resolution refutations. In this note, we will show an analogue for $Res^*(k)$ refutations.

A *literal* is a variable x or a negated variable $\bar{x}$. A *clause* is a disjunction $C = a_1 \vee \ldots \vee a_k$ of literals a_i, and a *term* is a conjunction $T = a_1 \wedge \ldots \wedge a_k$ of literals a_i. The *width* of a clause C or a term T is k, the number of literals in it. A k-term is a term of width $\ell \leq k$. We do not distinguish between a literal and a term consisting of one literal. Note that the negation $\bar{T}$ of a k-term T is a clause of width k, and vice versa.

A formula in *conjunctive normal form* (CNF) is a conjunction $F = C_1 \wedge \ldots \wedge C_m$ of clauses, and a formula in *disjunctive normal form* (DNF) is a disjunction $F = T_1 \vee \ldots \vee T_m$ of terms. A formula F in DNF is in k-DNF if every term T in F is a k-term.

The proof system of tree-like k-DNF Resolution, $Res^*(k)$ generalizes tree-like Resolution by allowing k-DNF formulas instead of just clauses as lines in the proof. The inference rules are (1) conjunction introduction, (2) k-cut and (3) weakening.

$$(1)\ \frac{F \vee T_1 \qquad G \vee T_2}{F \vee G \vee (T_1 \wedge T_2)} \qquad\qquad (2)\ \frac{F \vee T \qquad G \vee \bar{T}}{F \vee G} \qquad\qquad (3)\ \frac{F}{F \vee T}$$

where F and G are k-DNF formulas, T is a k-term and T_1 and T_2 are k-terms such that $T_1 \wedge T_2$ is still a k-term. Note that the resolution rule is a special case of (2) for $k = 1$.

A $Res^*(k)$-derivation of a k-DNF formula C from a set of k-DNF formulas F is an ordered binary tree, in which every node v is labeled with a k-DNF formula C_v such that:

1. The root is labeled with C.

2. If v an inner node with one predecessor u, then C_v follows from C_u by the weakening rule.

3. If v is an inner node with 2 predecessors u_1, u_2, then C_v follows from C_{u_1} and C_{u_2} by one of the inference rules conjunction introduction or k-cut.

4. If v is a leaf, then C_v is a formula from F.

A $Res^*(k)$-refutation of a CNF formula F is a $Res^*(k)$-derivation of the empty clause from F. Since the inference rules preserve satisfiability, a $Res^*(k)$-refutation of $Res^*(k)$-refutation of F shows that F is unsatisfiable. The proof system $Res^*(k)$ is also complete in the sense that a formula F has an $Res^*(k)$-refutation if and only if F is unsatisfiable, which follows from the completeness of tree-like resolution.

Let $Taut_k(F)$ be the set of tautologies $T \vee \bar{T}$ for all k-terms T over the variables of F. In some presentations, the proof system $Res^*(k)$ is defined as having these formulas as additional axioms. We show that they are redundant.

Theorem 1. *$F + Taut_k(F)$ has a $Res^*(k)$ refutation of size s iff F has a $Res^*(k)$ refutation of size at most s.*

We define a relation between terms in formulas occurring in a $Res^*(k)$-derivation. In a conjunction introduction inference (1), we call the term $T_1 \wedge T_2$ a child of the term T_1 in the left premise as well as of the term T_2 in the right premise. Moreover, each term in $F \vee G$ is a child of the same term appearing in F or in G in the premises. Analogously, in a cut inference (2), each term in $F \vee G$ is a child of the same term appearing in F or in G in the premises, and similarly for a weakening inference (3) and the terms appearing in F. The relation of a term being a descendant of another term in the proof is the transitive closure of this relation of being a child.

Proof. Let $T \vee \bar{T}$ be an axiom from $Taut_k(F)$, where $T = (a_1 \wedge \ldots \wedge a_k)$ is a k-term.

Consider the path from a leaf labelled $T \vee \bar{T}$ to the root. At some point on that path, a cut must be performed on a descendent of T, i.e., there is an inference of $F \vee G$ from $(T \wedge B) \vee F$ and $\bar{T} \vee \bar{B} \vee G$, where $(T \wedge B)$ is the descendent of T with $B = T_1 \wedge \ldots \wedge T_m$, the terms T_j being added to B in that order.

Let this path be labelled with the formulas $F_1, \ldots, F_n$ in that order, where $F_1 = T \vee \bar{T}$ and $F_n = (T \wedge B) \vee F$. Each line F_i is of the following form:

$$F_i \;=\; (T \wedge B_i) \vee A_1^{(i)} \vee \ldots \vee A_{\ell_i}^{(i)} \vee F_i'$$

where $(T \wedge B_i)$ is a descendent of T with $B_i = T_1 \wedge \ldots \wedge T_{j_i}$, the terms A_j are descendents of the literals $\bar{a}_j$ in F_1, and F_i' are all the remaining terms in F_i.

Now we replace each line F_i by the following line

$$F_i^* \;=\; B_i' \vee A_1^{(i)} \vee \ldots \vee A_{\ell_i}^{(i)} \vee F_i' \vee G$$

where $B_i' = \bar{T}_{j_i+1} \vee \ldots \vee \bar{T}_m$.

For every $i < n$, the formula F_{i+1} is derived from F_i by one of the following inferences:

1. a weakening, adding a term S to F'_i so that $F'_{i+1} = F'_i \vee S$,
2. a cut, which can only affect one or more terms in F'_i, or some of the terms A^i_j,
3. an $\wedge$-introduction increasing some term in F'_i or one of the terms A^i_j,
4. an $\wedge$-introduction adding a term T_{j_i+1} to B_i, where the second premise is $H \vee T_{j_i+1}$, so that $j_{i+1} = j_i + 1$, and $F'_{i+1} = F'_i \vee H$.

In the first case, the line F^*_{i+1} is obtained from F^*_i by the same weakening inference.

In cases 2 and 3, the formula F^*_{i+1} is similarly obtained from F^*_i by the same inference, using the same second premise which is obtained by the same derivation as in the original proof.

In the last case, we obtain F^*_{i+1} from F^*_i by a cut on $\bar{T}_{j_i+1}$ with the same second premise $H \vee T_{j_i+1}$, which is derived by the same derivation.

Note that the first line F^*_1 is $\bar{T} \vee \bar{B} \vee G$, for which we can use its original derivation, and that the last line F^*_m is $F \vee G$. Thus we have derived the conclusion $F \vee G$ by a smaller derivation, and we have eliminated this use of the axiom $T \vee \bar{T}$. Inductively, we can eliminate all uses of axioms from $Taut_k(F)$. □

References

[1] Juan Luis Esteban, Nicola Galesi, and Jochen Messner. On the complexity of resolution with bounded conjunctions. *Theoretical Computer Science*, 321(2-3):347–370, August 2004. Preliminary version in *ICALP '02*.

[2] Jan Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae*, 170(1-2):123–140, 2001.