Theoretical Computer Science 120 (1993) 157–168
Elsevier

Note

Cut formulas in propositional logic

Wenhui Zhang

Telemark College, N-3800 Bø i Telemark, Norway

Communicated by M. Nivat
Received October 1991
Revised October 1992

1. Introduction

Cut formulas in propositional logic can speed up some proofs exponentially [5] (a cut-free system is used as reference when we talk about speed-ups); it is hence important to study proof systems with cut, types of cut formulas (atomic cuts versus general cuts) and also relations between cut and techniques which may speed up proofs. In Section 2, we try to explain the importance of atomic cut formulas. We study the resolution principle [3] and analysis trees [4] with atomic cut and conjecture that there are proofs by analysis trees with atomic cuts and refutations by resolution, such that transferring them to cut-free proofs will cause exponential increase of the proof length. We also study unit resolution and conclude that transferring refutations by unit resolution to cut-free proofs does not cause exponential increase of the proof length. In Section 3, we discuss using definitions in analysis trees and in resolution. We conclude that it corresponds to using abbreviations in analysis trees and to adding possibilities to use more complicated cut formulas in resolution.

2. Atomic cuts

We discuss atomic cuts in propositional analysis trees, in resolution and in unit resolution. The following are some notations we are going to use in this section:

- $d \vdash_0 \Delta$: d is a proof of Δ by cut-free analysis trees;

Correspondence to: W. Zhang, Economic-Administr. Studies, Telemark Regional College, 3800 Bø i Telemark, Norway.

- $d \vdash_1 \Delta$: d is a proof of Δ by analysis trees with atomic cuts;
- $d \vdash \Delta \to F$: d is a resolution tree which deduces F from the set Δ;
- $\|d\|$: the number of nodes in the proof tree d.

2.1. *Analysis trees with atomic cuts*

We first present a proof system called analysis trees [4]. Let φ, ψ be formulas and Δ be a set of formulas; an analysis tree is a proof that uses the following rules:

- Normal rules:
  - A: $\Delta, \varphi, \neg\varphi$ if φ is atomic;
  - $\wedge: \dfrac{\Delta, \varphi \quad \Delta, \psi}{\Delta, \varphi \wedge \psi}$;
  - $\vee_0: \dfrac{\Delta, \varphi}{\Delta, \varphi \vee \psi} \qquad \vee_1: \dfrac{\Delta, \psi}{\Delta, \varphi \vee \psi}$.
- Atomic cut:

$$\frac{\Delta, \varphi \quad \Delta, \neg\varphi}{\Delta} \quad \text{if } \varphi \text{ is atomic.}$$

In this section, we only discuss atomic cuts. Contraction is implicitly used in the system, since the premisses and the conclusion of a rule are considered as sets of formulas.

The cut-elimination theorem [4] tells us that eliminating atomic cuts may cause the proof length to grow exponentially. We try to provide an example to match the upper bound of cut elimination. Let Γ be

$$\{\neg P_0, \neg Q_0, P_m, Q_m\} \cup \{(P_i \wedge Q_j) \wedge (\neg P_{i+1} \wedge \neg Q_{j+1}) \mid 0 \leqslant i \leqslant m-1, \; 0 \leqslant j \leqslant m-1\}.$$

We first prove Γ by an analysis tree with $O(m^2)$ steps and show that eliminating atomic cuts by some cut-elimination strategies leads to an exponential increase of the proof length, and conjecture that transferring the proof to cut-free proofs causes exponential increase of the proof length for all possible cut-elimination strategies.

A proof by an analysis tree with atomic cuts

Let Γ_j be $\{(P_i \wedge Q_j) \wedge (\neg P_{i+1} \wedge \neg Q_{j+1}) \mid 0 \leqslant i \leqslant m-1\}$. Γ_j is a subset of Γ.

Definition 2.1. We define T_{ij} $(1 \leqslant i \leqslant m,\ 0 \leqslant j \leqslant m-1)$ as parts of the analysis tree:

- $$T_{mj} =_{\text{def}} \dfrac{\dfrac{P_{m-1}, \neg P_{m-1} \quad \Gamma_j, Q_j, \neg Q_j}{\Gamma_j, \neg P_{m-1}, \neg Q_j, P_{m-1} \wedge Q_j} \qquad \dfrac{P_m, \neg P_m \quad Q_{j+1}, \neg Q_{j+1}}{\neg P_m \wedge \neg Q_{j+1}, P_m, Q_{j+1}}}{\Gamma_j, P_m, \neg P_{m-1}, \neg Q_j, Q_{j+1}}$$

- T_{ij} $(1 \leqslant i \leqslant m-1) =_{\text{def}}$

$$\dfrac{\dfrac{P_{i-1}, \neg P_{i-1} \quad Q_j, \neg Q_j}{P_{i-1} \wedge Q_j, \neg P_{i-1}, \neg Q_j} \qquad \dfrac{\begin{matrix} T_{i+1,j} \\ \Gamma_j, P_m, \neg P_i, \neg Q_j, Q_{j+1} \end{matrix} \quad Q_{j+1}, \neg Q_{j+1}}{\Gamma_j, P_m, \neg P_i \wedge \neg Q_{j+1}, \neg Q_j, Q_{j+1}}}{\Gamma_j, P_m, \neg P_{i-1}, \neg Q_j, Q_{j+1}}.$$

Proposition 2.2. *$T_{1j} \vdash_0 \Gamma_j, P_m, \neg P_0, \neg Q_j, Q_{j+1}$ and $\| T_{1j} \| = 6m+1$ for $j = 0, \ldots, m-1$.*

By the definition, we obtain $T_{ij} \vdash_0 \Gamma_j, P_m, \neg P_{i-1}, \neg Q_j, Q_{j+1}$ with $\| T_{ij} \| = 6(m+1-i)+1$. We obtain this proposition by replacing i with 1. In T_{ij}, both of the number of leaf nodes containing $Q_j, \neg Q_j$ and the number of leaf nodes containing $Q_{j+1}, \neg Q_{j+1}$ are $m-i+1$. This number will be useful in later discussions of cut elimination.

Definition 2.3. We define S_i $(0 \leqslant i \leqslant m-1)$ in terms of T_{1j} $(0 \leqslant j \leqslant m-1)$:
- $S_0 =_{\text{def}} T_{10}$,
- S_j $(1 \leqslant j \leqslant m-1) =_{\text{def}}$

$$\dfrac{\begin{matrix} S_{j-1} \\ \Gamma_0, \ldots, \Gamma_{j-1}, P_m, \neg P_0, \neg Q_0, Q_j \end{matrix} \qquad \begin{matrix} T_{1j} \\ \Gamma_j, P_m, \neg P_0, \neg Q_j, Q_{j+1} \end{matrix}}{\Gamma_0, \ldots, \Gamma_{j-1}, \Gamma_j, P_m, \neg P_0, \neg Q_0, Q_{j+1}}.$$

Proposition 2.4. *$S_{m-1} \vdash_1 \Gamma$ and $\| S_{m-1} \| = \mathrm{O}(m^2)$.*

By the definition, we obtain $S_j \vdash_1 \Gamma_0, \ldots, \Gamma_{j-1}, \Gamma_j, P_m, \neg P_0, \neg Q_0, Q_{j+1}$ with $\| S_j \| = 6mj + 6m + 2j + 1$. We obtain this proposition by replacing j with $m-1$.

Cut elimination

We define a cut-elimination strategy for atomic cuts.

Definition 2.5. Assume that d_0, d_1 and d_2 are cut-free. $t_1(d)$ is the result of transferring the proof tree d to the cut-free proof tree by the following strategy:

- $$\dfrac{\Pi_1,\varphi,\neg\varphi \quad \overset{d_0}{\Pi_0,\varphi}}{\Pi_1,\Pi_0,\varphi} \quad \text{reduces to} \quad \left(\begin{array}{c} d'_0 \\ \Pi_1,\Pi_0,\varphi \end{array}\right)$$

- $$\dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi}}{\Pi_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi_1,\Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi'_1,\Pi_0}}{\Pi_1,\Pi_0}$$

- $$\dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi} \quad \overset{d_2}{\Pi''_1,\varphi}}{\Pi_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi_1,\Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi'_1,\Pi_0} \quad \dfrac{\overset{d_2}{\Pi''_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi''_1,\Pi_0}}{\Pi_1,\Pi_0}$$

- $$\dfrac{\dfrac{\overset{d_1}{\Pi'_1} \quad \overset{d_2}{\Pi''_1,\varphi}}{\Pi_1,\varphi} \quad \overset{d_0}{\Pi_0,\varphi}}{\Pi_1,\Pi_0} \quad \text{reduces to} \quad \dfrac{\overset{d_1}{\Pi'_1} \quad \dfrac{\overset{d_2}{\Pi''_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi''_1,\Pi_0}}{\Pi_1,\Pi_0}$$

- $$\dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi} \quad \overset{d_2}{\Pi''_1}}{\Pi_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi_1,\Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_1}{\Pi'_1,\varphi} \quad \overset{d_0}{\Pi_0,\neg\varphi}}{\Pi'_1,\Pi_0} \quad \overset{d_2}{\Pi''_1}}{\Pi_1,\Pi_0}.$$

Lemma 2.6. *If $d_0 \vdash_0 \Pi_0, A$ and $d_1 \vdash_0 \Pi_1, \neg A$, and k is the number of leaf nodes of the form (or containing) $\{A, \neg A\}$ in the proof tree d_0, we can construct $d \vdash_0 \Pi_0, \Pi_1$ with $\|d\| = \|d_0\| + k\cdot(\|d_1\| - 1)$ by using the cut-elimination strategy t_1.*

The application of the strategy is as follows: (i) $t_1(S_0)$ is S_0; (ii) $t_1(S_j)$ (for $j = 1, \ldots, m-1$) is the result of replacing all nodes of the form $Q_j, \neg Q_j$ in $t_1(S_{j-1})$ by T_{1j} and replacing Q_j by $\{Q_{j+1}, \Gamma_j\}$ in all nodes below T_{1j} in $t_1(S_{j-1})$.

- The number of nodes of the form $Q_1, \neg Q_1$ in $t_1(S_0)$ is m. Hence, the number of nodes in $t_1(S_1)$ is $6m+1+(m*(6m+1)-m) = 6m^2+6m+1$.
- The number of nodes of the form $Q_2, \neg Q_2$ in $t_1(S_1)$ is m^2. Hence, the number of nodes in $t_1(S_2)$ is $6m^2+6m+1+(m^2*(6m+1)-m^2) = 6m^3+6m^2+6m+1$.
- Generally, the number of nodes of the form $Q_j, \neg Q_j$ in $t_1(S_{j-1})$ is m^j and the number of nodes in $t_1(S_j)$ is $6m+1+m*(6m+1)-m+\cdots+m^j*(6m+1)-m^j = 6m^{j+1}+6m^j+\cdots+6m^3+6m^2+6m+1$.

Proposition 2.7. $\| t_l(S_{m-1}) \| = O(m^m)$.

It shows that there is no k such that $\| t_l(d) \| = O(\| d \|^k)$ for every proof tree d with atomic cut formulas. An important aspect of this example is that in many subproofs implicit contractions (a kind of resource sharing) have been carried out at the same time as the conjunction rule is applied. It is important because a condition to take advantage of atomic cut formulas and obtain short proofs is that both of the eliminated literals come from different subtrees of the proof.

The strategy moves the right branch of proofs to the left branch. An alternative strategy is to move the left branch of proofs to the right branch.

Definition 2.8. Assume that d_0, d_1 and d_2 are cut-free. $t_r(d)$ is the result of transferring the proof tree d to the cut-free proof tree by the following strategy:

- $$\dfrac{\overset{d_0}{\Pi_0, \varphi} \quad \Pi_1, \varphi, \neg\varphi}{\Pi_1, \Pi_0, \varphi} \quad \text{reduces to} \quad \left(\begin{matrix} d_0' \\ \Pi_1, \Pi_0, \varphi \end{matrix}\right)$$

- $$\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \dfrac{\overset{d_1}{\Pi_1', \varphi}}{\Pi_1, \varphi}}{\Pi_1, \Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \overset{d_1}{\Pi_1', \varphi}}{\Pi_1', \Pi_0}}{\Pi_1, \Pi_0}$$

- $$\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \dfrac{\overset{d_1}{\Pi_1', \varphi} \quad \overset{d_2}{\Pi_1'', \varphi}}{\Pi_1, \varphi}}{\Pi_1, \Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \overset{d_1}{\Pi_1', \varphi}}{\Pi_1', \Pi_0} \quad \dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \overset{d_2}{\Pi_1'', \varphi}}{\Pi_1'', \Pi_0}}{\Pi_1, \Pi_0}$$

- $$\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \dfrac{\overset{d_1}{\Pi_1'} \quad \overset{d_2}{\Pi_1'', \varphi}}{\Pi_1, \varphi}}{\Pi_1, \Pi_0} \quad \text{reduces to} \quad \dfrac{\overset{d_1}{\Pi_1'} \quad \dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \overset{d_2}{\Pi_1'', \varphi}}{\Pi_1'', \Pi_0}}{\Pi_1, \Pi_0}$$

- $$\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \dfrac{\overset{d_1}{\Pi_1', \varphi} \quad \overset{d_2}{\Pi_1''}}{\Pi_1, \varphi}}{\Pi_1, \Pi_0} \quad \text{reduces to} \quad \dfrac{\dfrac{\overset{d_0}{\Pi_0, \neg\varphi} \quad \overset{d_1}{\Pi_1', \varphi}}{\Pi_1', \Pi_0} \quad \overset{d_2}{\Pi_1''}}{\Pi_1, \Pi_0}.$$

Lemma 2.9. *If $d_0 \vdash_0 \Pi_0, A$ and $d_1 \vdash_0 \Pi_1, \neg A$, and k is the number of leaf nodes of the form (or containing) $\{A, \neg A\}$ in the proof tree d_1, we can construct $d \vdash_0 \Pi_0, \Pi_1$ with $\| d \| = \| d_1 \| + k \cdot (\| d_0 \| - 1)$ by using the cut-elimination strategy t_r.*

The application of the strategy is as follows: (i) $t_r(S_0)$ is S_0; (ii) $t_r(S_j)$ (for $j = 1, \ldots, m-1$) is the result of replacing all nodes of the form $Q_j, \neg Q_j$ in T_{1j} by

$t_r(S_{j-1})$ and replacing Q_j by $\{Q_0, \Gamma_0, \Gamma_1, \ldots, \Gamma_{j-1}\}$ in all nodes below $t_r(S_{j-1})$ in T_{1j}.

- The number of nodes of the form $Q_1, \neg Q_1$ in T_{11} is m. Hence, the number of nodes in $t_r(S_1)$ is $6m+1+(m*(6m+1)-m)=6m^2+6m+1$.
- The number of nodes of the form $Q_2, \neg Q_2$ in T_{12} is m. Hence, the number of nodes in $t_r(S_2)$ is $6m+1+(m*(6m^2+6m+1)-m)=6m^3+6m^2+6m+1$.
- Generally, the number of nodes in $t_r(S_j)$ is $6m^{j+1}+6m^j+\cdots+6m^3+6m^2+6m+1$.

Proposition 2.10. $\|t_r(S_{m-1})\| = \mathrm{O}(m^m)$.

Both the strategies are deterministic and lead to an exponential increase of the proof length. By combining the reduction steps in these two strategies and removing the assumption in the definition of the reduction steps, we obtain a nondeterministic strategy. It is possible to produce a cut-free proof tree with the number of nodes less than $\mathrm{O}(m^m)$. But it seems that the increase will still be exponential. For instance, if we use a mixed strategy $\langle \mathrm{r}, \mathrm{l}, \mathrm{r}, \mathrm{l}, \mathrm{l}, \mathrm{r}, \mathrm{l}, \mathrm{l}, \mathrm{l}, \mathrm{r}, \ldots\rangle$ which eliminates the topmost cut formula by t_r, eliminates the second cut formula by t_l and so on, the numbers of nodes in the subtrees produced by the first steps of cut elimination are

- $6m+1+6m\cdot m$,
- $6m+1+6m\cdot m\cdot 2$,
- $6m+1+6m\cdot m\cdot(2m+1)$,
- $6m+1+6m\cdot m\cdot(2m+2)$,
- $6m+1+6m\cdot m\cdot(3m+2)$,
- $6m+1+6m\cdot m\cdot(3m^2+2m+1)$.

We obtain $6m+1+6m\cdot m\cdot\sum_{i=2}^{n}(i\cdot m^{i-2})$ as the number of nodes in the cut free proof tree if $m=n(n+1)/2$. This number is also exponential.

Conjecture 2.11. *There are analysis trees with atomic cuts such that transferring them to cut free proofs causes exponential increase of the proof length.*

2.2. Resolution

The resolution rule looks like a rule with atomic cut. In first-order logic, resolution may speed up proofs double exponentially [1]. We study it in propositional logic here. Formulas to be falsified by resolution are a conjunction (of a set) of clauses (a clause is a disjunction of literals). The resolution rule is

$$\frac{A \vee F \qquad \neg A \vee G}{F \vee G},$$

where A is atomic and F, G are clauses (could be empty). $F \vee G$ is called the resolvent of the rule. The input set of clauses is not satisfiable if the empty clause is deducible. We show that transferring a refutation to a cut-free proof may increase the proof length exponentially. We use the example of the previous subsection and construct a similar proof by resolution. Let Δ be the negated set of Γ of the previous section.

We obtain

$$\Delta = \{P_0, Q_0, \neg P_m, \neg Q_m\} \cup \{\neg P_i \vee \neg Q_j \vee P_{i+1} \vee Q_{j+1} \mid 0 \leqslant i \leqslant m-1,$$
$$0 \leqslant j \leqslant m-1\}.$$

We first falsify Δ by resolution with $O(m^2)$ steps and then conjecture that transferring the refutation to cut-free proofs causes exponential increase of the proof length.

Definition 2.12. We define T'_{ij} $(1 \leqslant i \leqslant m,\ 0 \leqslant j \leqslant m-1)$ as parts of the refutation tree:

- $T'_{mj} =_{\text{def}} \dfrac{\neg P_m \qquad \neg P_{m-1} \vee \neg Q_j \vee P_m \vee Q_{j+1}}{\neg P_{m-1} \vee \neg Q_j \vee Q_{j+1}}$,

- $T'_{ij}(1 \leqslant i \leqslant m-1) =_{\text{def}} \dfrac{\overset{T'_{i+1,j}}{\neg P_i \vee \neg Q_j \vee Q_{j+1}} \qquad \neg P_{i-1} \vee \neg Q_j \vee P_i \vee Q_{j+1}}{\neg P_{i-1} \vee \neg Q_j \vee Q_{j+1}}$.

Proposition 2.13. $T'_{1j} \vdash \Delta \to \neg P_0 \vee \neg Q_j \vee Q_{j+1}$ *and* $\|T'_{1j}\| = 2m+1$.

By the definition, we obtain $T'_{ij} \vdash \Delta \to \neg P_{i-1} \vee \neg Q_j \vee Q_{j+1}$ with $\|T'_{ij}\| = 2(m-i+1)+1$. We obtain this proposition by replacing i with 1.

Definition 2.14. We define S'_i $(0 \leqslant i \leqslant m-1)$ in terms of T'_{1j} $(0 \leqslant j \leqslant m-1)$:

- $S'_0 =_{\text{def}} T'_{10}$;

- S'_j (for $j = 1, \ldots, m-1$) $=_{\text{def}} \dfrac{\overset{S'_{j-1}}{\neg P_0 \vee \neg Q_0 \vee Q_j} \qquad \overset{T'_{1j}}{\neg P_0 \vee \neg Q_j \vee Q_{j+1}}}{\neg P_0 \vee \neg Q_0 \vee Q_{j+1}}$.

Proposition 2.15. $S'_{m-1} \vdash \Delta \to \neg P_0 \vee \neg Q_0 \vee Q_m$ *and* $\|S'_{m-1}\| = O(m^2)$.

By the definition, we obtain $S'_j \vdash \Delta \to \neg P_0 \vee \neg Q_0 \vee Q_{j+1}$ with $\|S'_j\| = 2mj + 2j + 2m + 1$. We obtain this proposition by replacing j with $m-1$.

Since $P_0, Q_0, \neg Q_m$ are in Δ, the rest of the refutation is of constant length. To eliminate the cut formulas, we first transfer this refutation to a proof of Γ (which is the negation of Δ) by an analysis tree with atomic cut formulas and then perform cut elimination on it. Recall that $T'_{ij} \vdash \Delta \to \neg P_{i-1} \vee \neg Q_j \vee Q_{j+1}$ and $T_{ij} \vdash_0 \Gamma', \neg P_{i-1}, \neg Q_j, Q_{j+1}$ (Γ' is some subset of Γ). T'_{ij} corresponds to T_{ij}, since the applications of the resolution rule can be transferred to some applications of the conjunction rule of analysis trees. S'_j corresponds also to S_j (for transferring procedure from resolution to analysis trees, the reader may refer to [6]). On the basis of this transformation and the discussion in the previous subsection, we state the following conjecture.

Conjecture 2.16. *There are refutations by resolution such that transferring them to cut-free proofs causes exponential increase of the proof length.*

Contraction in resolution is also important. In the example (we used $\neg P_i \vee \neg Q_k \vee Q_{k+1}$ and $\neg P_{i-1} \vee \neg Q_k \vee P_i \vee Q_{k+1}$ to deduce $\neg P_{i-1} \vee \neg Q_k \vee Q_{k+1}$) an implicit contraction had been carried out at the same time as the resolution rule is applied. Q_{k+1} and $\neg Q_k$ which were eliminated in later steps in the resolution tree come from both of their premisses.

2.3. Unit resolution

Here we consider a restricted resolution called unit resolution. It restricts the resolution such that one of the parent clauses of a resolvent has to be a unit clause (i.e. a literal). In first-order logic, unit resolution may speed up some proof exponentially [6]. We study it in propositional logic in this section. To compare proofs by unit resolution with cut-free proofs, we need a cut-free proof method as reference. Cut-free analysis trees could be used, but we construct a special cut-free proof system in order to make the comparison easier. The system contains the following rules:

- Ax: $\Gamma, A, \neg A \vdash$ if A is atomic;
- R: $\dfrac{\Gamma, A \vdash \qquad \Gamma, \psi \vdash}{\Gamma, A \vee \psi \vdash}$ if A is a literal.

We can assume that in unit resolution only unit clauses can be used more than once, because if a clause other than a unit clause is used more than once, we can find a shorter refutation. Since all clauses other than a unit clause will only be used once, we can arrange the literals in a clause in such an order that they are to be removed (resolved with a unit clause) in the same order. We compare unit resolution and cut-free proofs by transferring a refutation by unit resolution to a proof by the cut-free system.

Proposition 2.17. *If there is a refutation of Γ by unit resolution with k steps, we can find a cut-free d such that d is a proof of $\Gamma \vdash$ and $\|d\| = O(k)$.*

This proposition is justified by the following strategy of transferring a proof by unit resolution to a proof by the two rules Ax and R.

- A unit resolution step

$$\dfrac{\begin{matrix} d_1 \\ A \end{matrix} \qquad \begin{matrix} d_0 \\ \neg A \vee A_1 \vee \cdots \vee A_k \end{matrix}}{A_1 \vee \cdots \vee A_k}$$

- transfers to

$$\dfrac{\Gamma', A, \neg A \vdash \qquad \Gamma', A, A_1 \vee \cdots \vee A_k \vdash}{\Gamma', A, \neg A \vee A_1 \vee \cdots \vee A_k \vdash}$$

- Γ' contains the set of the original formulas and the deduced formulas. d_1 and d_0 were used to produce A and $\neg A \vee A_1 \vee \cdots \vee A_k$ in the conclusion of the second proof tree, if they are not already in Γ'. The proof continues from the right branch.
- The last step in the resolution which deduces the empty clause corresponds to an axiom.

It shows that a step in unit resolution corresponds to using an axiom in cut-free proofs and the order of proof length by unit resolution corresponds to that of a cut-free proof. In unit resolution there is no contraction of literals in resolvents, since one of their parent clauses is a unit clause. Eliminating atomic cuts in such proofs will not duplicate subproofs.

3. Definitions in proofs

It has been shown that (i) the length of refutations of the pigeonhole example by resolution is exponential and, by extending resolution with definitions, the order of the proof length can be reduced to polynomial [2] and (ii) by extending the Frege system with definitions, we cannot reduce proof length very much [1]. In this section, we discuss using definitions in resolution and analysis trees with general cuts and the relation between definition and cut.

3.1. Definitions in analysis trees

Using new symbols which do not occur in the conclusion (the last formula of a proof) does not have any advantage with respect to proof length. In that case, the new symbols must be eliminated by applications of the cut rule. Introducing definitions is different. Let Δ be the original set of formulas to be proved. Let Γ be the set of definitions to be used in the proof. We first consider this using definitions in a proof of Δ as a proof of $\Gamma \rightarrow \Delta$ (i.e. the union of Δ and the negated set of Γ). In this case, the new symbols appear in the conclusion.

To obtain Δ, we apply a theory which says that adding definitions of new symbols does not affect the validity of the original statement, in order to remove the definitions represented by Γ from the conclusion $\Gamma \rightarrow \Delta$. This is not a step of a proof by an analysis tree. To transfer such a proof to an ordinary proof by an analysis tree, we first replace all defined symbols with their definition. There remain two problems after the replacement: (1) Some of the leaf nodes may be $\Pi, F, \neg F$, where F is not atomic. (2) We must prove Γ' (which is Γ with the defined symbols replaced by their definition), and use the cut rule to eliminate Γ' from $\Gamma' \rightarrow \Delta$.

If we only accept $\Pi, F, \neg F$ as an axiom if F is atomic, we need to add some trivial proofs of formulas of the form $F, \neg F$. But if we do not require F to be atomic, we need only to prove Γ', and the length of the proof corresponds to the number of definitions in Γ. We summarize our discussion as follows: introducing definitions corresponds to using generalized axioms (of the form $\Gamma, F, \neg F$, without restriction on F) and it also corresponds to using abbreviations. The conclusion is the same if we allow proofs with definitions in a less formal way, i.e. we allow substituting A for B without using any proof rules, if the symbol A is defined as the formula B.

Proposition 3.1. *Using definitions in analysis trees corresponds to providing possibilities to use abbreviations.*

Although using abbreviations cannot reduce the proof length very much, if the proof system allows generalized axioms, but it can make a formula shorter, and can therefore be important with respect to mechanical proof checks. Sometimes, if we remove intermediate definitions, the number of symbols in a definition can grow exponentially.

3.2. Definitions in resolution

We first present the pigeonhole example. The pigeonhole principle can be understood as that there is no injective mapping from a set with $n+1$ elements to a set with n elements [1]. We use P_{ij} to represent that the ith element in the first set maps to the jth element in the second set.

Let Γ_n be the set of formulas $\{P_{i1} \vee P_{i2} \vee \cdots \vee P_{in} \mid i=1,\ldots,n+1\}$ and Δ_n be the set of formulas $\{P_{ik} \wedge P_{jk} \mid k=1,\ldots,n \text{ and } 1 \leqslant i<j \leqslant n+1\}$. The pigeonhole principle can then be represented by $\Gamma_n \rightarrow \Delta_n$.

By resolution, a proof of $\Gamma_n \rightarrow \Delta_n$ is a derivation of the empty clause from the union of Γ_n and the negated set of Δ_n. We want to deduce the empty clause from the following set of formulas:

$\{P_{i1} \vee P_{i2} \vee \cdots \vee P_{in} \mid i=1,\ldots,n+1\}$ and

$\{\neg P_{ik} \vee \neg P_{jk} \mid k=1,\ldots,n \text{ and } 1 \leqslant i<j \leqslant n+1\}$.

Let us call the union of these two sets for $\mathscr{P}_n$. The idea is to derive $\mathscr{P}_{n-1}$:

$\{P'_{i1} \vee P'_{i2} \vee \cdots \vee P'_{i,n-1} \mid i=1,\ldots,n\}$ and

$\{\neg P'_{ik} \vee \neg P'_{jk} \mid k=1,\ldots,n-1 \text{ and } 1 \leqslant i<j \leqslant n\}$

from $\mathscr{P}_n$ by resolution with polynomial length.

To succeed, we need appropriate definitions to connect these two sets of symbols together. The set of definitions needed is as follows: $P'_{ij} \Leftrightarrow P_{ij} \vee (P_{i,n} \wedge P_{n+1,j})$ for $i=1,\ldots,n-1$ and $j=1,\ldots,n$ [1].

The same strategy can be applied to $\mathscr{P}_{n-1}, \mathscr{P}_{n-2}$ until $\mathscr{P}_1$ is deduced. The empty clause can be deduced from $\mathscr{P}_1$ by constant length.

Compared with the Frege systems in which using definitions does not affect the number of lines very much (but may affect the number of symbols used) [1], using definitions in resolution is more important with respect to the number of proof lines.

Consider again the pigeonhole example. The actual formulas involved in the refutation by extended resolution are the clauses in $\mathscr{P}_n$ and the clauses which represent the definitions. The latter can be removed without affecting the validity. If we replace all new symbols with their definition, the proof is still sound. But the proof is not an ordinary proof by resolution, because the resolution rule is applied to eliminating formulas with various length. In fact, the number of symbols in the definitions grows exponentially. We conclude with the following proposition.

Proposition 3.2. *Using definitions in resolution corresponds to providing possibilities to use more complicated cut formulas.*

In the discussion about using definitions in resolution, if Γ is the set of definitions and Δ is the original set of the clauses, resolution has to be carried out on the set $\Gamma \cup \Delta$. It is very complicated with informal substitutions, since if A is defined as $A_1 \vee A_2$, then we have to combine two clauses to produce $\neg A$ which is $\neg A_1 \wedge \neg A_2$.

4. Summary

We provided a proof of a formula such that eliminating atomic cuts in the proof by some deterministic cut-elimination strategies leads to an exponential increase of the proof length, and we conjecture that transferring the proof to cut-free proofs causes exponential increase of the proof length for all possible cut-elimination strategies. The result applies also to general resolution. Unit resolution does not have much advantage over cut-free systems. A step in unit resolution corresponds to using an axiom in a cut-free proof.

We also discussed the role of allowing definitions in analysis trees and in resolution. We conclude that allowing definitions in an analysis tree corresponds to extending the analysis tree with generalized axioms of the form $\Gamma, A, \neg A$ without restriction on A, and allowing definitions in resolution corresponds to providing possibilities to use more complicated cut formulas. Allowing definitions means more to resolution than to analysis trees, because resolution has only limited possibilities to use cut.

We tried to explain the importance of atomic cuts. It needs more research in order to confirm or falsify the conjecture. It also needs more research in order to clarify the relation between atomic cuts and general cuts. From a study in first-order logic [7] we can (with some refinements) conclude that in propositional logic, eliminating general cut formulas may cause double exponential increase of the proof length. The cut-elimination theorem only states that if a formula is first proved with cut formulas, and if we eliminate them by cut-elimination strategies, the proof length may increase. Since every propositional formula can be proved by exponential length, there cannot be any real double exponential speed-up. Whether there is a proof of some propositional formula with cut such that all possible cut-elimination strategies will lead to a double exponential increase of the proof length is an open question.

From the viewpoint of mechanical proof search, there are problems with cuts. As a referee has pointed out, introducing cuts makes a method more efficient because proofs may be smaller in some cases, but it also makes it less efficient because it is much more important how to apply the rules in each situation. The balance between these two aspects needs to be studied in order to develop more efficient proof methods with cuts.

Acknowledgment

This work was partly carried out at the Department of Linguistics and Philosophy, University of Oslo. I thank Herman R. Jervell for discussions and support to this work.

References

[1] S.A. Cook and R.A. Reckhow, The relative efficiency of propositional proof systems, *Symbolic Logic* **44** (1979) 36–50.
[2] A. Haken, The intractability of resolution, *Theoret. Comput. Sci.* **39** (1985) 297–308.
[3] J.A. Robinson, A machine oriented logic based on the resolution principle, *J. ACM* **12** (1965) 23–41.
[4] H. Schwichtenberg, Proof theory: some applications of cut-elimination, in: J. Barwise, ed., *Handbook of Mathematical Logic* (North-Holland, Amsterdam, 1977).
[5] R. Statman, Bounds for proof-search and speed-up in the predicate logic, *Ann. Math. Logic* **15** (1978) 225–287.
[6] W. Zhang, Cut elimination and automatic proof procedures, Research Report No. 112, Institute of Informatics, University of Oslo, 1988.
[7] W. Zhang, Cut elimination and automatic proof procedures, *Theoret. Comput. Sci.* **91** (1991) 265–284.