The Deduction Rule and Linear and Near-linear Proof Simulations

Maria Luisa Bonet*
Department of Mathematics
University of California, San Diego

Samuel R. Buss*
Department of Mathematics
University of California, San Diego

Abstract

We introduce new proof systems for propositional logic, *simple deduction Frege systems*, *general deduction Frege systems* and *nested deduction Frege systems*, which augment Frege systems with variants of the deduction rule. We give upper bounds on the lengths of proofs in Frege proof systems compared to lengths in these new systems. As applications we give near-linear simulations of the propositional Gentzen sequent calculus and the natural deduction calculus by Frege proofs. The length of a proof is the number of lines (or formulas) in the proof.

A general deduction Frege proof system provides at most quadratic speedup over Frege proof systems. A nested deduction Frege proof system provides at most a nearly linear speedup over Frege system where by "nearly linear" is meant the ratio of proof lengths is $O(\alpha(n))$ where α is the inverse Ackermann function. A nested deduction Frege system can linearly simulate the propositional sequent calculus, the tree-like general deduction Frege calculus, and the natural deduction calculus. Hence a Frege proof system can simulate all those proof systems with proof lengths bounded by $O(n \cdot \alpha(n))$. Also we show

*Supported in part by NSF Grant DMS-8902480.

that a Frege proof of n lines can be transformed into a tree-like Frege proof of $O(n \log n)$ lines and of height $O(\log n)$. As a corollary of this fact we can prove that natural deduction and sequent calculus tree-like systems simulate Frege systems with proof lengths bounded by $O(n \log n)$.

1 Introduction

A Frege proof system is an inference system for propositional logic in which the only rule of inference is modus ponens. Although it suffices to have modus ponens as the single inference rule to obtain a complete proof system, it is well-known that other modes of inference are also sound. A notable example of this is the deduction rule which states that if a formula B has a proof from an additional, extra-logical hypothesis A (in symbols, $A \vdash B$) then there is a proof of $A \supset B$. This paper considers various strengthenings of this deduction rule and establishes upper bounds on the proof-speedups obtained with these deduction rules.

By a "speedup" of a proof, we mean the amount that proofs can be shortened with additional inference rules. In this paper, the *length* or *size* of a proof is the number of lines in the proof; where a line consists of either a formula or a sequent (depending on the proof system). We write $\vdash_{\overline{k}} B$ (and $A_1, \ldots, A_s \vdash_{\overline{k}} B$) to indicate that the formula B has Frege proof of $\leq k$ lines (from the hypotheses $A_1, \ldots, A_s$). More generally, we write "$\vdash^{T}_{\overline{k}}$" to mean "provable in proof system T with $\leq k$ lines". If S and T are proof systems we say that S can linearly (respectively, quadraticly) simulate T if, for any T-proof of k lines, there is an S-proof of the same (or sometimes an equivalent[†]) formula of $O(k)$ lines (respectively, of $O(k^2)$ lines). We say that T provides at most linear (respectively, quadratic) speedup over S if S can linearly (respectively, quadraticly) simulate T. In general, we define:

Definition We say S *simulates* T with an increase in size of $f(x)$, if for any

[†]It is necessary to use an equivalent formula instead of the same formula in the case where S and T have different languages. In this paper, we shall only consider simulations between systems which have the same languages (i.e., the same logical connectives); see Reckhow [14] for an in-depth treatment of the more general case.

T-proof of k lines, there is an S-proof of the same formula of $O(f(k))$ lines. We say that T provides an at most $f(x)$ *speedup* if S can simulate T with an increase in size of $f(x)$.

An alternative, commonly used measure of the length of a propositional proof is the number of symbols in the proof. This is the approach used, for instance, by Cook-Reckhow [7, 14] and Statman [15]. They have also considered *extended Frege* proof systems which consist of Frege proof systems plus a new inference rule, called the *extension rule* which allows introduction of abbreviations for formulas. They have proved that the minimum number of lines in a Frege proof of a formula (plus the number of symbols in the formula proved) is linearly related to the minimum number of symbols in an extended Frege proof. The intuition behind their proof of this fact is that, by introducing abbreviations, it is possible to make all formulas in a proof very short. The linear relation between number of lines in a Frege proof and number of symbols in an extended Frege proof means that upper and lower bounds on the number of lines in Frege proofs translate into bounds on the number of symbols in extended Frege proofs, and vice-versa.

We begin by defining the main propositional proof systems used in this paper. The logical connectives of all our systems are presumed to be $\neg$, $\vee$, $\wedge$ and $\supset$; however, our results hold for any complete set of connectives.

Definition A *Frege proof system* (denoted $\mathcal{F}$) is characterized by:

(1) A finite set of axiom schemata. A axiom schema consists of a tautology and specifies that all instances of the tautology are axioms. For example, a possible axiom schema is $(A \supset (B \supset A))$; A and B represent arbitrary formulas.

(2) The only rule of inference is Modus Ponens:

$$\frac{A \qquad A \supset B}{B}$$

(3) A proof in this system is a sequence of formulas $A_1, \ldots, A_n$ (also called 'lines') where each A_i is either a substitution instance of an axiom schema or is inferred by Modus Ponens from some A_j and A_k with $j, k < i$. Such a sequence is, by definition, a proof of A_n.

(4) The proof system must be consistent and complete.

The *length* of an $\mathcal{F}$-proof is the number of lines in the proof; we write $\vdash_{\overline{k}} A$ to indicate that A has a Frege proof of length $\leq k$. We further write $A_1, \ldots, A_s \vdash_{\overline{k}} B$ to mean that B is provable from the hypotheses A_i with a Frege proof of $\leq k$ lines; in other words, that there is a sequence of $\leq k$ formulas, each of which is one of the A_i's, is an axiom, or is inferred by modus ponens from earlier formulas, such that B is the final formula of the proof. Although we have not specified the axiom schemata to be used in a Frege proof system, it is easy to see that different choices of axiom schemata will change the lengths of proofs only linearly. To see this, suppose $\mathcal{F}_1$ and $\mathcal{F}_2$ are Frege proof systems with different axiom schemata. Since $\mathcal{F}_2$ is complete, it can prove (every instance of) every axiom schema of $\mathcal{F}_1$; furthermore, if $\mathcal{F}_2$ proves axiom schema T of $\mathcal{F}_1$ in c lines, then $\mathcal{F}_2$ proves any instance of T in c lines (since instances of $\mathcal{F}_2$ axioms are still $\mathcal{F}_2$ axioms). Hence, since the Frege system $\mathcal{F}_1$ has a finite number of axiom schemata, there is a constant upper bound c' on the length of $\mathcal{F}_2$-proofs of axioms of $\mathcal{F}_1$. This allows any $\mathcal{F}_1$ proof of n lines to be transformed to an $\mathcal{F}_2$-proof of $\leq c'n$ lines by just replacing the $\mathcal{F}_1$ axioms by their $\mathcal{F}_2$ proofs.

The simplest form of the deduction theorem states that if $A \vdash B$ then $\vdash A \supset B$. This can be informally phrased as a rule in the form

$$\frac{A \vdash B}{A \supset B}$$

which is called the *1-simple deduction rule*; more generally, the *simple deduction rule* is

$$\frac{A_1, \ldots, A_n \vdash B}{A_1 \supset (\cdots (A_{n-1} \supset (A_n \supset B)) \cdots)}$$

We next define extensions to the Frege proof system that incorporate the deduction theorem as a rule of inference. For this purpose, the systems defined below have proofs in which the lines are *sequents* of the form $\Gamma \vDash A$; intuitively, the sequent means that the formulas in Γ tautologically imply A: operationally, a sequent $\Gamma \vDash A$ means that A has been proved using the formulas in Γ as assumptions.

A *general deduction Frege system* (denoted $d\mathcal{F}$) incorporates a strong version of the deduction rule. Each line in a general deduction Frege proof is a sequent of the form $\Gamma \models A$ where A is a formula and Γ is a set of formulas. When Γ is empty, we write just $\models A$. A general deduction Frege system is specified by a finite set of axiom schemata, which must be suitable for a Frege proof system. The four valid inference rules in a general deduction Frege proof are:

$$\models A \qquad - A \text{ an instance of an axiom schema}$$

$$\{A\} \models A \qquad - \text{Hypothesis}$$

$$\frac{\Gamma_1 \models A \supset B \qquad \Gamma_2 \models A}{\Gamma_1 \cup \Gamma_2 \models B} \qquad - \text{Modus Ponens}$$

$$\frac{\Gamma \models B}{\Gamma \setminus \{A\} \models A \supset B} \qquad - \text{Deduction Rule}$$

We write $A_1, \ldots, A_n \vdash^{d\mathcal{F}}_{k} B$ to indicate that $\{A_1, \ldots, A_n\} \models B$ has a general deduction Frege proof of $\leq k$ lines.

Deduction Frege systems are quite general since they allow hypotheses to be "opened" and "closed" (i.e., "assumed" and "discharged") in arbitrary order. A more restrictive system is the *nested deduction Frege* proof system which requires the hypotheses to be used in a 'nested' fashion. The nested deduction Frege systems are quite natural since they correspond to the way mathematicians actually reason while carrying out proofs. A second reason the nested deduction Frege system seems quite natural is that we shall prove below that nested deduction Frege proof systems can simulate with linear size proofs the propositional Gentzen sequent calculus, the tree-like general deduction Frege proofs and the natural deduction calculus.

The primary feature of the nested deduction Frege proof system is that hypotheses must be closed in reverse order of their opening. And after a hypothesis is closed, any formula proved inside the scope of the hypothesis is no longer available. The following is an example of a nested deduction Frege proof in a system in which $X \supset X$ and $X \supset (Y \supset (X \land Y))$ are among the axiom schemata.

$$
\begin{array}{rl}
\langle A\rangle \vDash A & \text{Hypothesis } A \text{ opened}\\
\langle A,B\rangle \vDash B & \text{Hypothesis } B \text{ opened}\\
\langle A,B\rangle \vDash A\supset A & \text{Axiom}\\
\langle A,B\rangle \vDash A & \text{Modus Ponens}\\
\langle A\rangle \vDash B\supset A & \text{Deduction Rule; } B \text{ closed}\\
\langle A,C\rangle \vDash C & \text{Hypothesis } C \text{ opened}\\
\langle A,C\rangle \vDash A\supset A & \text{Axiom}\\
\langle A,C\rangle \vDash A & \text{Modus Ponens}\\
\langle A\rangle \vDash C\supset A & \text{Deduction Rule; } C \text{ closed}\\
\langle A\rangle \vDash (B\supset A)\supset((C\supset A)\supset((B\supset A)\wedge(C\supset A))) & \text{Axiom}\\
\langle A\rangle \vDash (C\supset A)\supset((B\supset A)\wedge(C\supset A)) & \text{Modus Ponens}\\
\langle A\rangle \vDash (B\supset A)\wedge(C\supset A) & \text{Modus Ponens}\\
\langle\rangle \vDash A\supset((B\supset A)\wedge(C\supset A)) & \text{Deduction Rule; } A \text{ closed}
\end{array}
$$

A sequent in a nested deduction Frege ($nd\mathcal{F}$) proof is of the form $\Gamma \vDash A$ where now Γ is a *sequence* of formulas. An $nd\mathcal{F}$ proof is a sequence of sequents $\Gamma_i \vDash A_i$ ($i = 1, 2, \ldots, n$) such that Γ_0 is taken to be the empty sequence and, for each i, one of the following holds:

(a) $\Gamma_i = \Gamma_{i-1}$ and A_i is an axiom.

(b) $\Gamma_i = \Gamma_{i-1} * \langle A_i \rangle$. This opens an assumption, $*$ denotes concatenation of sequences.

(c) Γ_{i-1} is $\Gamma_i * \langle B \rangle$ and A_i is $B \supset A_{i-1}$. This is the deduction rule.

(d) $\Gamma_i = \Gamma_{i-1}$ and A_i is inferred from A_j and A_k by Modus Ponens where each of $\Gamma_j \vDash A_j$ and $\Gamma_k \vDash A_k$ are available to sequent i. We say sequent j is available to sequent i if $j < i$ and for all ℓ, if $j < \ell < i$ then Γ_j is an initial subsequence of Γ_ℓ.[‡]

Examples of Modus Ponens as defined in (d) can be seen in proof above. Note that there are two sequents containing $A \supset A$; this is since the first one is not available to the second Modus Ponens. Interestingly, for this particular

[‡]It is also possible, though less elegant, to define sequent j being available to sequent i iff $j < i$ and Γ_j is an initial subsequence of Γ_{i-1}. Our Main Theorems still apply with this alternative definition.

example, the above proof could be shortened one line by removing the two sequents containing $A \supset A$ and adding either $\langle\rangle \models A \supset A$ or $\langle A\rangle \models A \supset A$ as the first or second line of the proof (respectively). In general however, nested deduction Frege proofs might have to repeat lines if they depend on different hypotheses.

We write $A_1, \ldots, A_n \vdash^{nd\mathcal{F}}_{k} B$ if $\langle A_1, \ldots, A_n\rangle \models B$ has a nested deduction Frege proof with $\leq k$ sequents.

Nested deduction Frege proofs can be conveniently represented in pictorial form as a column of formulas with vertical bars that represent the opening, closing and availability of assumptions. This is best defined by an example; the $nd\mathcal{F}$-proof given above would be pictorially represented as:

$$\begin{array}{l}
\left[\begin{array}{l}
A \\
\left[\begin{array}{l} B \\ A \supset A \\ A \end{array}\right. \\
B \supset A \\
\left[\begin{array}{l} C \\ A \supset A \\ A \end{array}\right. \\
C \supset A \\
(B \supset A) \supset ((C \supset A) \supset ((B \supset A) \wedge (C \supset A))) \\
(C \supset A) \supset ((B \supset A) \wedge (C \supset A)) \\
((B \supset A) \wedge (C \supset A))
\end{array}\right. \\
A \supset ((B \supset A) \wedge (C \supset A))
\end{array}$$

Nested deduction Frege proofs are conceptually simple and natural and, in practice, seem to simplify the process of discovering proofs. Thus, it is surprising that a Frege proof system can simulate nested deduction Frege proofs with near-linear size proofs: this fact is the content of our main theorems below.

This paper also uses the propositional sequent calculus and the propositional natural deduction calculus.

The results of this paper consist of various simulation results between proof systems. Many of these simulations are linear; e.g., section 4.2 shows that nested deduction Frege proofs linearly simulate natural deduction proofs and sequent calculus proofs. But some of our simulations are near-linear or quadratic; this includes the following results (and others): Frege systems simulate nested deduction Frege systems, natural deduction and the sequent calculus with an increase in size of $O(n\alpha(n))$ where α is the inverse Ackermann function; tree-like Frege proofs can simulate non-tree-like Frege proofs with an increase in size of $O(n \log n)$; and Frege proof systems can quadraticly simulate general deduction Frege proof systems. It remains an open problem whether these non-linear simulations can be improved; indeed it is possible that all the proof systems considered in this paper linearly simulate each other. There appear to be fundamental difficulties in proving that our simulations are optimal, because the most likely methods of proving optimality would involve establishing non-linear lower bounds on the lengths of proofs in Frege proof systems (which are the weakest proof systems we discuss). However, to the best of present-day knowledge, it is entirely possible that all tautologies have linear size Frege proofs. That is to say, it is open whether all tautologies containing n logical connectives have Frege proofs of $O(n)$ lines.

This last question is interesting because of close connections to open problems in computational complexity. Cook-Reckhow [7] noted that if there is *any* proof system (which must have polynomial time recognizable proofs) in which tautologies have proofs containing polynomially many symbols, then $NP = coNP$. From this it follows immediately that if tautologies always have Frege proofs with polynomially many lines, then $NP = coNP$. This is because Frege proofs with polynomially many lines can be transformed into extended Frege proofs with polynomially many symbols. Furthermore, because of the connections (due to Cook) between extended Frege proof systems and theories of bounded arithmetic, superpolynomial lower bounds on the number of lines in Frege proofs imply that S^1_2 does not prove $P = NP$ [6, 5]. Thus it is expected to be quite difficult to give superpolynomial lower bounds on the lengths of Frege proofs; in fact, it already appears to be quite difficult to give non-linear lower bounds.

The problem of giving non-linear lower bounds on the number of lines in Frege proofs is similar in spirit to the notoriously difficult problem of giving non-linear lower bounds on the size of circuits for computing explicitly given Boolean functions. However, we feel that the former problem may be more amenable to solution than the latter. The best that we have been able to achieve in the way of lower bounds is to show that our method of proof for obtaining the $O(n\alpha(n))$ upper bound can not be improved (see section 3.2); of course, this does not rule out alternative methods of proof yielding $O(n)$ size proofs.

The present paper is an expansion of portions of [1, 3]. The proof of the main theorem of the present paper depends on the Serial Transitive Closure problem for trees; for this, see [4] (preliminary versions of this are also in [1, 3]).

2 Simulations of Simple and General Deduction

This section contains some preliminary results giving bounds on how much the deduction rule can shorten proofs. Most of our results are stated in the form "If proof system X can prove a formula in n lines, then the formula has a Frege proof of $f(n)$ lines". Obviously f depends on the system X.

First recall the usual proof of the deduction theorem (see e.g. Kleene [8]) which establishes the following:

Theorem 1 *(Deduction Theorem) There is a constant c such that if $A \vdash_{\overline{n}} B$ then $\vdash_{\overline{c \cdot n}} A \supset B$.*

(The constant c is equal to 5 in Kleene's system). The bound of $c \cdot n$ is obtained by replacing each formula C occurring in a proof of B from A with the formula $A \supset C$ and then "filling in the gaps" in the resulting proof with a constant number of lines per gap. For axioms, this is easily done since if C is an axiom then $A \supset C$ can be proved in a constant number of lines. For modus ponens inferences, this is done using the fact that for any formulas A, C and D there is a Frege proof of $A \supset D$ from $A \supset C$ and $A \supset (C \supset D)$ with a constant number of lines. If the proof of Theorem 1 is iterated for m hypotheses, then we get the result that if $A_1, \ldots, A_m \vdash_{\overline{n}} B$ then $\vdash_{\overline{c^m n}} A_1 \supset (A_2 \supset \cdots \supset (A_m \supset B) \cdots)$. However we can substantially improve the bound $c^m n$:

Theorem 2 *(Simple Deduction Theorem) Suppose $A_1, \ldots, A_m \vdash_{\overline{n}} B$. Then*

$$\vdash_{\overline{O(n+m)}} (A_1 \supset (A_2 \supset \cdots \supset (A_m \supset B) \cdots)).$$

Proof Given an n line proof P of B from assumptions $A_1, \ldots, A_m$, we construct a Frege proof P' of B from the single assumption $A_1 \wedge A_2 \wedge \cdots \wedge A_m$ (where the conjunction is to be associated from left-to-right). For any Frege proof system, there is a constant k such that $C \wedge D \vdash_{\overline{k}} C$ and $C \wedge D \vdash_{\overline{k}} D$. Thus, P' can be constructed to (1) first derive each of $A_1, \ldots, A_m$ in $2k(m-1)$ lines and (2) then derive B in $\leq n$ lines (via P). Clearly P' has $O(m+n)$ lines and by one application of Theorem 1, there is a Frege proof of $A_1 \wedge \cdots \wedge A_m \supset B$ with $O(m+n)$ lines. Finally it can be shown by induction on m that

$$\vdash_{\overline{O(m)}} \quad [A_1 \wedge \cdots \wedge A_m \supset B] \supset \\ [(A_1 \supset (A_2 \supset \cdots \supset (A_m \supset B) \cdots))].$$

By combining these last two proofs with a Modus Ponens inference, Theorem 2 is proved. $\square$

We use the name *simple deduction Frege* proof system for the system in which all hypotheses must be opened at the beginning a proof and closed at the end of a proof. Theorem 2 shows that a Frege proof system can simulate simple deduction Frege proofs with linear size proofs; or equivalently, that the simple deduction Frege proof system provides only a *linear speedup* (i.e., a *constant factor speedup*) over Frege proof systems.

An interesting corollary to Theorem 1 is that conjunctions may be arbitrarily reordered and reassociated with linear size Frege proofs:

Corollary 3 *Let B be any conjunction of $A_1, \ldots, A_m$ in that order but associated arbitrarily. Let $i_1, \ldots, i_n$ be any sequence from $\{1, \ldots, m\}$ and let C be any conjunction of $A_{i_1}, \ldots, A_{i_n}$ again in the indicated order and associated arbitrarily. Then*

$$\vdash_{\overline{O(m+n)}} B \supset C.$$

Proof By Theorem 1 it suffices to show that $B \vdash_{\overline{O(m+n)}} C$. The proof of B from C proceeds as follows: (1) from the assumption B deduce each subformula of B and, in particular, each of the formulas $A_1, \ldots, A_m$, and (2) deduce each subformula of C from the smallest to the largest. Since there is a constant k such that $E \wedge F \mathrel{\vdash_{\overline{k}}} E$ and $E \wedge F \mathrel{\vdash_{\overline{k}}} F$ and $E, F \mathrel{\vdash_{\overline{k}}} E \wedge F$ for all formulas E and F, it is clear that the proof contains $O(m+n)$ lines. $\square$

We now consider the simulation of proof systems with more powerful versions of the deduction rule.

Theorem 4 *If* $\mathrel{\vdash^{d\mathcal{F}}_{n}} B$ *then* $\mathrel{\vdash_{\overline{O(n^2)}}} B$.

Theorem 4 states that a general deduction Frege proof system can provide no more than a quadratic speedup over a Frege proof system; whether this quadratic bound is optimal is an open question.

Proof For the proof, we let $\bigwedge\limits_{i=1}^{m} A_i$ denote any conjunction of the formulas A_i ordered and associated arbitrarily (each A_i should occur exactly once as a conjunct). To prove Theorem 4, we prove the more general result that if $\{A_1, \ldots, A_m\} \vDash B$ has a $d\mathcal{F}$-proof P of n lines then $(\bigwedge\limits_{i=1}^{m} A_i) \supset B$ has a Frege proof P' of $O(n^2)$ lines. To form the proof P' replace each sequent $\{A_1, \ldots, A_m\} \vDash B$ of P by the formula $(\bigwedge\limits_{i=1}^{m} A_i) \supset B$; it will suffice to "fill in the gaps" to make P' a valid proof. First, an axiom in P becomes w.l.o.g. an axiom of the Frege system. Second, a hypothesis $\{A\} \vDash A$ in P becomes the tautology $A \supset A$ which has a constant length Frege proof.

Third, the sequents in a Modus Ponens inference in P

$$\frac{\Gamma_1 \vDash A \supset B \qquad \Gamma_2 \vDash A}{\Gamma_1 \cup \Gamma_2 \vDash B}$$

become the formulas $\bigwedge \Gamma_1 \supset (A \supset B)$ and $\bigwedge \Gamma_2 \supset A$ and $\bigwedge(\Gamma_1 \cup \Gamma_2) \supset B$. It will suffice to show that the third formula can be proved from the first formulas with a Frege proof of $O(n)$ lines. By Corollary 3 there are Frege proofs of $\bigwedge(\Gamma_1 \cup \Gamma_2) \supset \Gamma_i$ for $i = 1, 2$ containing $O(m)$ lines where $\Gamma_1 \cup \Gamma_2$ contains m formulas. From these latter two formulas and from $\bigwedge \Gamma_1 \supset (A \supset B)$ and $\bigwedge \Gamma_2 \supset B$ there is a Frege proof of $\bigwedge(\Gamma_1 \cup \Gamma_2) \supset B$ with a constant number of lines. It is easily shown that the number of formulas in the lefthand side of sequent in a $d\mathcal{F}$-proof is bounded by the number of lines in the proof; hence $m \leq n$ and for Modus Ponens, one can "fill in the gap" in P' with $O(n)$ lines.

Fourth, the sequents in a deduction rule inference in P

$$\frac{\Gamma_1 \vDash B}{\Gamma_2 \vDash A \supset B}$$

where Γ_2 is $\Gamma_1 \setminus \{A\}$ become the formulas $\bigwedge\Gamma_1 \supset B$ and $\bigwedge\Gamma_2 \supset (A \supset B)$. In this case, by Corollary 3, there is a Frege proof of $(A \wedge \bigwedge\Gamma_2) \supset \bigwedge\Gamma_1$ of $O(n)$ lines (again since the number of formulas in the conjunction is bounded by n). With this, there is a Frege proof of $\bigwedge\Gamma_2 \supset (A \supset B)$ from $\bigwedge\Gamma_1 \supset B$ with constantly many additional lines. Thus we have "filled the gap" for the deduction rule inference with $O(n)$ lines. $\square$

3 Simulation of the Nested Deduction Frege System

In this section we show how to almost linearly simulate the nested deduction Frege system by the Frege system. In section 3.1 we reduce the simulation problem to the problem of solving the serial transitive closure problem on trees, and in section 3.2 we discuss a combinatorial result about the serial transitive closure problem.

3.1 Main Theorems

We next state our main results that Frege systems can simulate nested deduction Frege proof systems with nearly linear proof size. The "near-linear" size estimates are in terms of extremely slow growing functions such as log^* and the inverse Ackermann function. The log^* function is defined so that $\log^* n$ is equal to the least number of iterations of the logarithm base 2 which applied to n yield a value < 2. In other words, $\log^* n$ is equal to the least value of k such that $n < 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}$ where there are k 2's in the stack. To get even slower growing functions, we define the $\log^{(*i)}$ functions for each $i \geq 0$. The $log^{(*0)}$ function is just the base 2 logarithm function and the $log^{(*1)}$ is just the log^* function. For $i > 1$, the $log^{(*i)}$ function is defined to be equal to the least number of iterations of the $log^{(*i-1)}$ function which applied to n yields a value < 2. The Ackermann function can be defined by the equations:

$$
\begin{aligned}
A(0,m) &= 2m \\
A(n+1,0) &= 1 \\
A(n+1,m+1) &= A(n,A(n+1,m))
\end{aligned}
$$

It can be shown that $A(i+1,j)$ is equal to the least value n such that $\log^{(*i)}(n) \geq j$, this means that $\log^{(*i)} A(i+1,j) = j$ (see [4] for a proof of this fact). It is well-known that the Ackermann function is recursive but dominates eventually every primitive recursive function.

Definition The *inverse Ackermann* function α is defined so that $\alpha(n)$ is equal to the least value of i such that $A(i,i) > n$. Equivalently, $\alpha(n)$ is equal to the least i such that $\log^{(*i-1)} n < i$.

Main Theorem 5 *Let $i \geq 0$. Suppose $\vdash^{nd\mathcal{F}}_{n} B$ and that in this $nd\mathcal{F}$-proof of B assumptions are opened m times. Then*

$$\vdash_{O(n+m\log^{(*i)} m)} B.$$

Main Theorem 6 *If $\vdash^{nd\mathcal{F}}_{n} B$ then $\vdash_{O(n\cdot\alpha(n))} B$.*

These Main Theorems are extremely close to a linear simulation of nested deduction Frege proof systems by Frege proof systems. It is immediate that $m < n$; Theorem 5 implies that if one could somehow further bound the number of hypotheses m by $O(n/\log^{(*i)} n)$ for a fixed value i, then one would obtain a linear simulation. However, we have no indication that m can be bounded in this way.

To prove the Main Theorems, we reduce them to combinatorial theorems regarding the serial transitive closure of trees. Suppose that we have a nested deduction Frege proof P of a sequent $\vDash B$ such that P contains n lines and uses the hypothesis rule m times. To prove Main Theorem 5 for a fixed value of i, we will translate P into a Frege proof of B containing $O(n+m\log^{(*i)} m)$ lines. Likewise, for Main Theorem 6, P is translated into a Frege proof of B of $O(n\cdot\alpha(n))$ lines.

This is how the simulation goes: Each line in the proof P is of the form $\Gamma \vDash B$ where Γ is a sequence of formulas $\langle A_1, \ldots, A_k \rangle$. From the sequent $\Gamma \vDash B$ we form the logically equivalent formula $(\bigwedge \Gamma) \supset B$ where the conjunction is associated from left to right; thus $\bigwedge \Gamma$ is the formula $((\cdots(A_1 \wedge A_2) \wedge \cdots \wedge A_{k-1}) \wedge A_k)$. When Γ is empty, $\bigwedge \Gamma$ is a fixed tautology. This translation of sequents into equivalent formulas gives us a sequence of formulas P'; unfortunately, P' is not a valid Frege proof and so it remains to show how P' can be made into a valid Frege proof with only a relatively small increase in the number of lines.

To make P' into a valid Frege proof we shall add additional lines. There are four rules of inference for $nd\mathcal{F}$: Axiom, Hypothesis, Deduction Rule and Modus Ponens. For each rule of inference, we explain what lines need to be added to P'; we save Modus Ponens for last since it is by far the most difficult case.

First consider an axiom inference in P which is of the form $\Gamma \vDash B$ where B is an axiom, so P' contains $\bigwedge \Gamma \supset B$ as the corresponding formula. Since a Frege proof system is axiomatized with axiom schemata, there is a proof of $B \supset (X \supset B)$ with a constant number of lines (independent of the formulas B and X). Thus there is a constant length Frege proof of the formula $\bigwedge \Gamma \supset B$; namely, take the axiom B, derive $B \supset (\bigwedge \Gamma \supset B)$ and then use modus ponens. This constant length Frege proof is inserted into the sequence P'.

Second, consider a hypothesis inference where P contains a sequent $\Gamma * \langle B \rangle \vDash B$ and P' contains $((\bigwedge \Gamma) \wedge B) \supset B$. It is easy to derive $(X \wedge B) \supset B$ in a Frege proof in a constant number of lines where the constant is independent of the formulas B and X, so the sequent in P' can be derived in a constant number of lines.

Third, consider a deduction rule inference in P; here P contains a sequent $\Gamma * \langle A \rangle \vDash B$ followed immediately by $\Gamma \vDash A \supset B$ and P' contains the corresponding formulas $((\bigwedge \Gamma) \wedge A) \supset B$ and $(\bigwedge \Gamma) \supset (A \supset B)$. Again there is a constant length Frege proof of the latter formula in P' from the former one; this constant length proof is to be inserted into P'.

Fourth and hardest, we consider a line in P' that corresponds to a sequent of P obtained by Modus Ponens. Suppose that in P there are lines $\Gamma_1 \vDash A$

and $\Gamma_2 \vDash A \supset B$ from which $\Gamma \vDash B$ is inferred by Modus Ponens. Since P is a nested deduction Frege proof, Γ_1 and Γ_2 are initial subsequences of Γ. In P' the formulas $(\bigwedge \Gamma_1) \supset A$ and $(\bigwedge \Gamma_2) \supset (A \supset B)$ appear and from them we wish to derive the formula $(\bigwedge \Gamma) \supset B$ in a small number of lines. Note that there is a constant size Frege proof of $X \supset B$ from the hypotheses $X_1 \supset A$ and $X_2 \supset (A \supset B)$ and $X \supset X_1$ and $X \supset X_2$ where the constant is independent of the formulas X, X_1, X_2, A and B. Thus we will modify P' by adding the formulas $(\bigwedge \Gamma) \supset (\bigwedge \Gamma_i)$ at the beginning and inserting a Frege proof of $(\bigwedge \Gamma) \supset B$ from these new formulas and from the other two formulas.

It remains now to give short Frege proofs of the formulas $(\bigwedge \Gamma) \supset (\bigwedge \Gamma')$ which have been added to the beginning of P'. By examining the fourth case above we see that there are $< 2n$ such formulas and they always have Γ' an initial subsequence of Γ and thus they are tautologies of the form

$$((\cdots(A_1 \wedge A_2) \wedge \cdots) \wedge A_k) \supset ((\cdots(A_1 \wedge A_2) \wedge \cdots) \wedge A_\ell)$$

where without loss of generality $\ell < k \leq m$. To prove one such tautology requires $O(k-\ell)$ lines. Unfortunately, if we used $O(k-\ell)$ lines for each tautology, the total number of lines would only be bounded by $O(m \cdot n)$ instead of the desired bound of $O(n + m \log^{(*i)} m)$ or $O(n \cdot \alpha(n))$. To get this lower bound on the number of lines we must exploit the fact that there are many tautologies to be proved. In other words, we can achieve significant reduction in the number of proof lines by proving the $2n$ many tautologies simultaneously rather than separately.

To get short Frege proofs, we shall rephrase the problem as a transitive closure problem. We shall now work only with tautologies of the form $\bigwedge \Gamma \supset \bigwedge \Pi$ where Π is a proper initial subsequence of Γ and may be the empty sequence. Since there were m uses of the hypothesis rule in P, there are at most $m+1$ distinct $\bigwedge \Gamma$'s; we think of them forming a directed graph G (actually a tree) with an edge from $\bigwedge \Gamma$ to $\bigwedge \Pi$ iff Γ extends Π by one element. For example, for the nested deduction proof pictured in section 1, the directed graph G of tautologies is:

$$\begin{tikzcd} & \bigwedge\langle\rangle & \\ & \bigwedge\langle A\rangle \arrow[u] & \\ \bigwedge\langle A,B\rangle \arrow[ur] & & \bigwedge\langle A,C\rangle \arrow[ul] \end{tikzcd}$$

There are $\leq 2n$ distinct "target" tautologies which are, by definition, the ones we need to prove. It is useful to think of these target tautologies as being in the transitive closure of the directed graph G. The Frege proof of these target tautologies will proceed as follows: First prove in $O(m)$ lines the tautologies $\bigwedge\Gamma \supset \bigwedge\Pi$ where Γ extends Π by a single element (this may include both target and non-target tautologies), obtaining all the edges in the tree-like directed graph G. Next we prove all the target tautologies in $O(n + m \log^{(*i)} m)$ or $O(n \cdot \alpha(n))$ lines. The procedure for this latter step is to prove many intermediate formulas $\bigwedge\Gamma \supset \bigwedge\Pi$ from the transitive closure of the directed graph of $\bigwedge\Gamma$'s. For this, we consider the slightly more general setting of the Serial Transitive Closure problem discussed next.

3.2 Serial Transitive Closure Problem

A directed graph is transitive if, whenever there is an edge from a node X to a node Y and an edge from Y to Z, then there is an edge from X to Z. The transitive closure of G is a smallest transitive, directed graph containing G. We write $X \to Y$ to indicate the presence of an edge from X to Y. It is easy to see that any edge in the transitive closure of a graph G can be obtained from the edges of G by a series of zero or more *closure steps*, which are inferences of the form

$$\frac{A \to B \qquad B \to C}{A \to C}$$

In other words, if $A \to B$ and $B \to C$ are edges in the transitive closure of G, then $A \to C$ is too. This is because the edges that can be derived by closure steps from edges in G both must be in any transitive graph containing G and also form a transitive graph on the nodes of G.

The serial transitive closure problem is the problem of deriving a given set of "closure edges" in the transitive closure of a directed graph. A solution to the serial transitive closure problem is a sequence of closure steps which generates all of the given closure edges and the size of a solution is the number of closure steps in the solution. The serial transitive closure problem is formally defined as follows:

Serial Transitive Closure Problem:

An *instance* consists of

- A directed graph G with m edges and
- A list of n *closure* edges $X_i \rightarrow Y_i$ $(i = 1, \ldots, n)$ which are in the transitive closure of G but not in G.

A *solution* is a sequence of edges $U_i \rightarrow V_i$ $(i = 1, \ldots, s)$ containing all n closure edges such that each $U_i \rightarrow V_i$ is inferred by a single closure step from earlier edges and/or edges in G. We call s the number of steps of the solution.

Note that the number of steps in a solution counts only closure steps and does not count edges that are already in G.

It should be stressed that the set of closure edges can be any subset of the edges in the transitive closure of the graph (but not in the graph). The degenerate case of deriving a single closure edge $A \rightarrow B$ is quite simple, since the minimum number of closure steps required will be one less than the length of a shortest path from A to B. The general question of determining the optimal size of a solution is made difficult by the fact that, when a *set* of closure edges is being derived, it may be possible for individual closure steps to aid in the generation of multiple closure edges. In other words, it is not necessary to generate each closure edge independently. It is also important that the set of closure edges will, in general, not be *all* the edges in the transitive closure of the graph; the problem of finding a minimal length derivation of all the edges in the transitive closure of the graph is uninteresting because, in this case, exactly one closure step is needed per closure edge.

We call a directed graph a *tree* if it can be obtained from a rooted tree T by either directing all edges in T away from the root or directing all edges towards the root. We picture trees as having root at the top and either having all edges directed downwards or having all edges directed upwards.

Theorem 7 *Let $i \geq 0$. If the directed graph G is a tree then the serial transitive closure problem has a solution with $O(n + m \log^{(*i)} m)$ steps.*

Theorem 8 *If the directed graph G is a tree then the serial transitive closure problem has a solution with $O((n+m) \cdot \alpha(m))$ steps.*

Theorems 7 and 8 are precisely what is needed to complete the proof of the Main Theorems. This is because in section 3.1 the proof of the Main Theorems was reduced to the problem of proving $\leq 2n$ 'target' tautologies. Let G be the graph with m edges defined at the end of section 3.1 and let the target tautologies be the closure edges: this specifies an instance of the Serial Transitive Closure problem and any solution of this instance leads directly to a Frege proof of the target tautologies of length $O(s)$; namely, the Frege proof simulates each closure step in the solution with a constant number of lines.

Unfortunately the proofs of Theorems 7 and 8 are too complicated to include in this paper. Complete proofs can be found in [1, 3, 4]. In addition, [4] proves a lower bound of $O(n\alpha(n))$ for the serial transitive closure problem, showing that Theorem 8 is optimal. Of course, this does not rule out other approaches for obtaining Frege proof system simulations of the nested deduction Frege proof system; so it is still open whether the Frege calculus can linearly simulate $nd\mathcal{F}$.

4 Applications

We discuss and prove some corollaries which give unexpected connections between nested deduction Frege proof systems and tree-like $d\mathcal{F}$-proofs, natural deduction and the propositional Gentzen sequent calculus.

4.1 Simulation of the Tree-like General Deduction Frege System

A proof is *tree-like* if no line is used more than once in the proof as a hypothesis of an inference.

Theorem 9 *If $\Gamma \vDash A$ has a tree-like general deduction Frege proof of n lines, then $\vdash^{nd\mathcal{F}}_{O(n)} \Gamma_{\text{seq}} \vDash A$ where Γ_{seq} is any sequence containing the same elements as the set Γ without repetition.*

Note One elementary fact to note about $nd\mathcal{F}$-proofs is that if Π is a sequence of k formulas and if the sequent $\Pi \vDash A$ has an $nd\mathcal{F}$-proof of n lines, then the sequent also has an $nd\mathcal{F}$-proof of n lines in which the first k lines are hypothesis inferences which open the hypotheses in Π— of course these k hypotheses remain open at the end of the proof. By reordering the first k lines of the $nd\mathcal{F}$-proof, it is clear that for any permutation Π' of Π, $\Pi' \vDash A$ also has an n line $nd\mathcal{F}$-proof.

Notation The subscript *seq* will be omitted most of the time. When we deal with a nested deduction Frege proof, we will often abuse notation by writing $\Gamma \models A$ instead of $\Gamma_{\text{seq}} \models A$. By the previous note the order of the formulas in Γ_{seq} is irrelevant.

Proof of the theorem. We shall prove by induction on n that, if the sequent $\{A_1, \ldots, A_m\} \models B$ has a tree-like $d\mathcal{F}$-proof P of n lines then there is an $nd\mathcal{F}$-proof P' of $\langle A_1, \ldots, A_m \rangle \vDash B$ of length $\leq 2n$ lines. The proof splits into four cases depending on the final inference in P.

Case 1: The last line of P is $\models A$, for A an axiom. Then P' is just an $nd\mathcal{F}$-proof of $\models A$ which has one line.

Case 2: The last line of P is $\{A\} \models A$. Then P' is an $nd\mathcal{F}$-proof of $\langle A \rangle \models A$ which has only one line.

Case 3: The last line of P is

$$\frac{\Gamma_1 \models A \supset B \qquad \Gamma_2 \models A}{\Gamma_1 \cup \Gamma_2 \models B}$$

Assume the proof of $\Gamma_1 \models A \supset B$ has n_1 lines and the proof of $\Gamma_2 \models A$ has n_2 lines, so that $n = n_1 + n_2 + 1$ since P is tree-like. By the induction hypothesis, there are $nd\mathcal{F}$-proofs P_1 and P_2 of the sequents $\Gamma_1 \models A \supset B$ and $\Gamma_2 \models A$ of lengths $\leq 2n_1$ and $\leq 2n_2$ lines, respectively. The proof P' of $(\Gamma_1 \cup \Gamma_2)_{\text{seq}} \models B$ is:

$$\begin{array}{|ll}
\Gamma_1 \cup \Gamma_2 & \\
\left.\begin{array}{c}\vdots\\ A \supset B\end{array}\right\} & \text{from } P_1 \\
\left.\begin{array}{c}\vdots\\ A\end{array}\right\} & \text{from } P_2 \\
B & \text{by modus ponens}
\end{array}$$

This proof has $\leq 2n_1 + 2n_2 + 1$ lines, i.e., $\leq 2n$ lines.

Note The first line $\lceil \Gamma_1 \cup \Gamma_2$ above means that each formula in $\Gamma_1 \cup \Gamma_2$ is opened as a hypothesis.

Case 4: The last line of P is:

$$\frac{\Gamma \models C}{\Gamma \setminus \{A\} \models A \supset C}$$

By the induction hypothesis, there is a $nd\mathcal{F}$-proof P_1 of the sequent $\Gamma \models C$ with $2n-2$ lines . The proof of $(\Gamma \setminus \{A\})_{\text{seq}} \models A \supset C$ is:

$$\begin{array}{|l}
\Gamma \setminus \{A\} \\
\left.\begin{array}{|c} A \\ \vdots \\ C \end{array}\right\} \text{ from } P_1 \\
A \supset C
\end{array}$$

This proof has size $(2n-2)+1$ or $(2n-2)+2$ lines, depending on whether A is in Γ or not. In either case, this is $\leq 2n$ lines.

The theorem follows from cases 1–4. $\square$

Corollary 10 *If A has a tree-like $d\mathcal{F}$-proof of n lines, then* $\vdash_{\overline{O(n\cdot\alpha(n))}} A$.

4.2 Simulation of the Sequent Calculus

The next theorems give a linear simulation of the propositional Gentzen sequent calculus by the nested deduction Frege system. We will prove this by showing the stronger result that the nested deduction Frege system linearly simulates the Gentzen calculus where we do not count structural inferences like weakening, contraction and exchange. There is a direct simulation of the sequent calculus, but we prove the stronger result because we need it for the simulation of the natural deduction system. For the next theorems, it is crucial that Gentzen sequent calculus proofs are always tree-like. The definition of the Gentzen sequent calculus can be found in Takeuti [17]; we are concerned with only the propositional fragment of the sequent calculus, which we call PKT. Also we work with a version PKT* of the propositional sequent calculus where we do not count the weak structural inferences weakening, exchange and contraction. In other words, the size of a PKT*-proof is computed by counting only sequents which are inferred by an inference other than these three kinds of weak structural rules. Because we use PKT*, Theorems 11 and 14 and Corollary 15 also hold for many variations of the sequent calculus, for instance with the mix rule, or with a rule that allows arbitrary reordering of cedents, or with either the multiplicative or additive versions of rules. (But the tree-like property is crucial for our proofs.)

Recall that $*$ denotes concatenation of sequences. So if $\Gamma = \langle A_1, \ldots, A_l \rangle$ and $\Delta = \langle B_1, \ldots, B_s \rangle$ then $\Gamma * \Delta = \langle A_1, \ldots, A_l, B_1, \ldots, B_s \rangle$ Also, if $\Delta = \langle B_1, \ldots, B_s \rangle$ then $\neg\Delta = \langle \neg B_1, \ldots, \neg B_s \rangle$

Theorem 11 *If $A_1, \ldots, A_m \longrightarrow B_1, \ldots, B_k$ has a PKT proof of n steps, then* $\vdash^{nd\mathcal{F}}_{O(n)} \langle A_1, \ldots, A_m, \neg B_1, \ldots, \neg B_k \rangle \models p \wedge \neg p$

Proof We prove, by induction on n, the following stronger statement:

> If $\vdash^{\mathrm{PKT}^*}_{n} A_1, \ldots, A_m \longrightarrow B_1, \ldots, B_k$, then there is a subsequence Π of $\langle A_1, \ldots, A_m, \neg B_1, \ldots, \neg B_k \rangle$ such that $\Pi \vdash^{nd\mathcal{F}}_{O(n)} p \wedge \neg p$.

Theorem 11 follows directly from this fact since: (a) if Π is a subsequence of $\langle A_1, \ldots, A_m, \neg B_1, \ldots, \neg B_k \rangle$ and $\Pi \vdash^{nd\mathcal{F}}_{s} p \wedge \neg p$, then $\vdash^{nd\mathcal{F}}_{s+m+k} \langle A_1, \ldots, A_m, \neg B_1, \ldots, \neg B_k \rangle \models p \wedge \neg p$, and (b) it is easy to prove that if a PKT sequent calculus proof has n lines then every sequent in the proof has no more than n lines in the antecedent or in the succendent (i.e., that $m, k \leq n$).

Let P be a tree-like PKT^* proof of n inferences other than weak structural inferences; P' will be the $nd\mathcal{F}$ proof that we are going to create to simulate P.

Base Case: $n = 1$. The proof P consists only of an axiom, say $A \longrightarrow A$. Now we need to build a $nd\mathcal{F}$ proof of $\langle A, \neg A \rangle \models p \wedge \neg p$. That can be done in constant number of lines, say d.

Induction Step: We suppose the statement holds for all $m < n$ and prove it for n. We have different cases depending on what the last line of the PKT*-proof is. We will prove the most representative cases:

$\neg$: ***left*** The last line of P is

$$\frac{\Gamma \longrightarrow \Delta, A}{\neg A, \Gamma \longrightarrow \Delta}$$

By the induction hypothesis, there is a $nd\mathcal{F}$-proof of $\Pi \models p \wedge \neg p$ where Π is a subsequence of $\Gamma * \neg \Delta * \langle \neg A \rangle$, in say $c(n-1)$ lines. There is an almost identical proof of $\Pi_1 \models p \wedge \neg p$ in the same number of lines, where Π_1 is a subsequence of $\langle \neg A \rangle * \Gamma * \neg \Delta$, i.e. a reordering of Π. The only difference between the two proofs is in the order of the hypotheses, which is unimportant (recall the note following Theorem 9).

$\vee$: ***right*** The last line of P is

$$\frac{\Gamma \longrightarrow \Delta, A}{\Gamma \longrightarrow \Delta, A \vee B} \quad (\text{or} \quad \frac{\Gamma \longrightarrow \Delta, B}{\Gamma \longrightarrow \Delta, A \vee B})$$

Since $\Gamma \longrightarrow \Delta, A$ has a proof of $n-1$ lines, by the induction hypothesis there is a ndF-proof, say P_1 of

$$\Pi \models p \wedge \neg p$$

in $c(n-1)$ lines, where Π is a subsequence of $\Gamma * \neg \Delta * \langle \neg A \rangle$. If $\neg A$ is not in the sequence Π, then the same proof in $c(n-1)$ lines is a proof of $p \wedge \neg p$ from a subsequence of $\Gamma * \neg\Delta * \langle\neg(A \vee B)\rangle$. If $\neg A$ is in Π then let Π_1 be obtained from Π by replacing $\neg A$ with $\neg(A \vee B)$. The following is a proof of $\Pi_1 \models p \wedge \neg p$:

$$\left[\begin{array}{ll} \Pi_1 & \\ \left.\begin{array}{c} \vdots \\ \neg A \end{array}\right\} & \begin{array}{l}\text{a proof of } \neg A \text{ from } \neg(A \vee B) \\ \text{in } c_1 \text{ lines.}\end{array} \\ \left.\begin{array}{c} \vdots \\ p \wedge \neg p \end{array}\right\} & \text{from } P_1 \end{array}\right.$$

All the lines above are obtained either from P_1 or in a constant number of lines, say c_1. So this proof has $\leq c(n-1)+c_1$ lines, so $\leq c \cdot n$ lines taking c such that $c_1 \leq c$.

$\vee$: ***left*** The last line of P is:

$$\frac{A, \Gamma \longrightarrow \Delta \qquad B, \Gamma \longrightarrow \Delta}{A \vee B, \Gamma \longrightarrow \Delta}$$

Say that $A, \Gamma \longrightarrow \Delta$ has a proof of n_1 lines, $B, \Gamma \longrightarrow \Delta$ has a proof of n_2 lines, so that $n = n_1 + n_2 + 1$, since the proofs of $A, \Gamma \longrightarrow \Delta$ and $B, \Gamma \longrightarrow \Delta$ do not share work (P is tree-like). By the induction hypothesis there are $nd\mathcal{F}$ proofs P_1 and P_2 of

$$\Pi_1 \models p \wedge \neg p \qquad \text{and} \qquad \Pi_2 \models p \wedge \neg p$$

of sizes $c \cdot n_1$ and $c \cdot n_2$ respectively, where Π_1 and Π_2 are subsequences of $\langle A\rangle * \Gamma * \neg\Delta$ and $\langle B\rangle * \Gamma * \neg\Delta$ respectively. If A is not in Π_1 (or B is not in Π_2), then the same proof P_1 (or P_2) is a proof of $p \wedge \neg p$ from the subsequence Π_1 (or Π_2) of $\langle A \vee B\rangle * \Gamma * \neg\Delta$. If A is in Π_1 and B is in Π_2, then consider the sequence Π_3 containing all the formulas from Π_1 except A, and all the formulas (non repeated) from Π_2 except B. The following is a proof of $\langle A \vee B\rangle * \Pi_3 \models p \wedge \neg p$:

$$
\left[\begin{array}{l}
A \vee B \\
\left[\begin{array}{l}
\Pi_3 \\
\left.\left[\begin{array}{c} A \\ \vdots \\ p \wedge \neg p \end{array}\right.\right\} \qquad \text{from } P_1 \\
A \supset p \wedge \neg p \\
\left.\left[\begin{array}{c} B \\ \vdots \\ p \wedge \neg p \end{array}\right.\right\} \qquad \text{from } P_2 \\
B \supset p \wedge \neg p \\
\qquad \vdots \\
A \vee B \supset p \wedge \neg p \\
p \wedge \neg p
\end{array}\right.
\end{array}\right.
$$

This proof has $c \cdot n_1 + c \cdot n_2 + c_2$ lines for some constant c_2. Since $n = n_1 + n_2 + 1$ the proof length is $\leq c \cdot n$ taking $c_2 \leq c$.

Note Even though P_1 and P_2 had the hypotheses in a different order in which they are used in the proof above, we can always obtain a proof in the same number of lines where the hypotheses can be permuted.

Cut The last line of P is

$$\frac{\Gamma \longrightarrow \Delta, A \qquad A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}$$

Suppose $\Gamma \longrightarrow \Delta, A$ has a PKT*-proof of n_1 lines, and $A, \Gamma \longrightarrow \Delta$ has a PKT*-proof of n_2 lines, and since P is tree-like, $n = n_1 + n_2 + 1$. By the induction hypothesis, there are proofs of $\Pi_1 \models p \wedge \neg p$ and $\Pi_2 \models p \wedge \neg p$, say P_1 and P_2 of $\leq c \cdot n_1$ and $\leq c \cdot n_2$ lines respectively, and where Π_1 and Π_2 are subsequences of $\Gamma * \neg\Delta * \langle \neg A\rangle$ and $\langle A\rangle * \Gamma * \neg\Delta$ respectively. Again, if $\neg A$ is not in Π_1 (or A is not in Π_2), then the same proof P_1 (or P_2) is a proof of $p \wedge \neg p$ from a subsequence of $\Gamma * \neg\Delta$ in $\leq c \cdot n$ lines.

But if $\neg A$ is in Π_1 and A is in Π_2, then define Π_3 as a sequence containing the formulas of Π_1 except $\neg A$, and the formulas of Π_2 except A. The following is a proof of $p \wedge \neg p$ from the subsequence Π_3 of $\Gamma * \Delta$.

$$
\left[\begin{array}{l}
\quad \Pi_3 \\
\left.\left[\begin{array}{c} \neg A \\ \vdots \\ p \wedge \neg p \end{array}\right.\right\} \qquad \text{from } P_1 \\
\neg A \supset p \wedge \neg p \\
\left.\left[\begin{array}{c} A \\ \vdots \\ p \wedge \neg p \end{array}\right.\right\} \qquad \text{from } P_2 \\
A \supset p \wedge \neg p \\
\qquad \vdots \\
A \vee \neg A \supset p \wedge \neg p \\
\qquad \vdots \\
p \wedge \neg p
\end{array}\right.
$$

This proof has $\leq c \cdot n_1 + c \cdot n_2 + c_3$ lines, for some constant c_3. So $\leq c \cdot n$ lines for $c \geq c_3$.

Weakening The last line of P is:

$$\frac{\Gamma \longrightarrow \Delta}{A, \Gamma \longrightarrow \Delta} \quad (\text{or} \quad \frac{\Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, A})$$

Recall that since P is a PKT*-proof, we are not counting the weakening inferences; so, by the induction hypothesis, there is a $nd\mathcal{F}$-proof say P_1 of

$$\Pi \models p \wedge \neg p$$

of $c \cdot n$ lines, where Π is a subsequence of $\Gamma * \neg\Delta$. Since Π is also a subsequence of $A * \Gamma * \neg\Delta$, P_1 is also a proof of $p \wedge \neg p$ from the sequence Π.

The result follows taking $c \geq d, c_1, c_2, c_3$. $\square$

Before we finish the result on the $nd\mathcal{F}$ simulation of sequent calculus, let us state the following lemmas:

Lemma 12 *The formula* $\neg(B_1 \vee \cdots \vee B_k) \supset (\neg B_1 \wedge \cdots \wedge \neg B_k)$ *has a proof in* $O(k)$ *lines in a Frege system. Here, parentheses are associated from left to right.*

Proof By induction on k. □

Lemma 13 *If* $\vdash^{nd\mathcal{F}}_{n} \langle A_1, \ldots, A_m, \neg B_1, \ldots, \neg B_k \rangle \models p \wedge \neg p$, *then*

$$\vdash^{nd\mathcal{F}}_{O(n)} \langle A_1, \ldots, A_m \rangle \models B_1 \vee \cdots \vee B_k.$$

Proof Let $\Gamma = \langle A_1, \ldots, A_m \rangle$. Assume that $\vdash^{nd\mathcal{F}}_{n} \Gamma * \langle \neg B_1, \ldots, \neg B_k \rangle \models p \wedge \neg p$. We build a $nd\mathcal{F}$-proof of $\Gamma \models B_1 \vee \cdots \vee B_k$ the following way:

$$
\begin{array}{l}
\Gamma \\
\left.\begin{array}{l}
\neg B_1 \wedge \cdots \wedge \neg B_k \\
\quad \vdots \\
\neg B_1 \\
\neg B_2 \wedge \cdots \wedge \neg B_k \\
\quad \vdots \\
\neg B_{k-1} \\
\neg B_k \\
\quad \vdots \\
p \wedge \neg p
\end{array}\right\} \begin{array}{l}\text{$O(k)$ lines to separate } \neg B_1 \wedge \cdots \wedge \neg B_k \\ \text{and n lines by assumption}\end{array} \\
(\neg B_1 \wedge \cdots \wedge \neg B_k) \supset p \wedge \neg p \\
\left.\begin{array}{l}
\quad \vdots \\
\neg(B_1 \vee \cdots \vee B_k) \supset (\neg B_1 \wedge \cdots \wedge \neg B_k)
\end{array}\right\} \text{$O(k)$ lines from lemma 12} \\
\quad \vdots \\
\neg(B_1 \vee \cdots \vee B_k) \supset p \wedge \neg p \\
\quad \vdots \\
B_1 \vee \cdots \vee B_k
\end{array}
$$

The proof above has $O(k)+n+d$ lines, for some constant d. Since $k \leq n$, the proof has $O(n)$ lines. □

Theorem 11 and Lemma 13 together show that nested deduction Frege systems simulate tree-like sequent calculus with a linear increase in the size of the proof.

Theorem 14 *If* $\vdash^{\mathrm{PKT}}_{n} A_1, \ldots, A_m \longrightarrow B_1, \ldots, B_k$, *then*

$$\vdash^{nd\mathcal{F}}_{O(n)} \langle A_1, \ldots, A_m \rangle \models B_1 \vee \cdots \vee B_k.$$

Corollary 15 *If* $\vdash^{\mathrm{PKT}}_{n} \longrightarrow A$, *then* $\vdash_{O(n \cdot \alpha(n))} A$.

Proof If $\vdash^{\mathrm{PKT}}_{n} \longrightarrow A$, then by Theorem 14, $\vdash^{nd\mathcal{F}}_{O(n)} A$, and by Theorem 6, $\vdash_{O(n \cdot \alpha(n))} A$. □

Corollary 15 improves a theorem of Orevkov [11] which implies that if $\longrightarrow A$ has a (tree-like) proof in the sequent calculus of n lines and height h, then $\vdash_{O(n \log h)} A$. Orevkov's theorem is stated for a proof system KGL which is a reformulation of the usual sequent calculus [10]. Although KGL proofs need not be tree-like, it appears that Gentzen proofs must be tree-like in order to be linearly translated into KGL. Orevkov, like us, does not need to count structural inferences.

4.3 Simulation of the Natural Deduction

The next results give a linear simulation of the propositional natural deduction calculus by the nested deduction Frege system. For the definition of natural deduction see [13, 16]; it is important to note that natural deduction proofs are tree-like. We call the propositional portion of it ND. First we claim that PKT* linearly simulates ND, and as a corollary we obtain that nested deduction Frege systems linearly simulate ND. This corollary could be obtained by a direct simulation of ND by $nd\mathcal{F}$, but doing it this way also relates sequent calculus with natural deduction.

Theorem 16 *If A has a ND-proof of n lines, then* $\vdash^{\mathrm{PKT}^*}_{O(n)} \longrightarrow A$.

Proof By induction on n. Prove that if A has a ND-proof of n lines using hypotheses $A_1, \ldots, A_k$, then $\vdash^{PKT^*}_{O(n)} A_1, \ldots, A_k \longrightarrow A$. The details are left to the reader. □

Corollary 17 *If A has a ND-proof of n lines, then* $\vdash^{nd\mathcal{F}}_{O(n)} A$.

Proof The result follows directly from Theorem 16 and the proof of Theorem 14. □

Corollary 18 *If A has a ND-proof of n lines, then* $\vdash_{O(n\alpha(n))} A$.

Proof The result follows directly from Corollary 17 and Main Theorem 6. □

5 Tree-like Frege Proofs

In this last section, we prove that the tree-like Frege calculus simulates the Frege calculus with an increase in size of $n \log n$. In fact, we show that a Frege proof of n lines can be transformed into a tree-like Frege proof of $O(n \log n)$ lines and of height $O(\log n)$. This result improves on theorems of Krajíček [9] and Pitassi-Beame-Impagliazzo [12] which say that a Frege calculus proof of n lines can be transformed into a tree-like Frege proof of $O(n^2)$ lines and of height $O(\log n)$. We need some definitions and lemmas:

Definition Let $A_1, \ldots, A_n$ be formulas with n a power of 2. The *Balanced Conjunction* $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ of $A_1, \ldots, A_n$ is defined inductively by:

- if $n = 1$ then $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ is just A_1.
- Otherwise $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ is $\left(\bigwedge\!\!\!\bigwedge_{i=1}^{n/2} A_i\right) \wedge \left(\bigwedge\!\!\!\bigwedge_{i=1}^{n/2} A_{(n/2)+i}\right)$.

Definition Let $A_1, \ldots, A_n$ be formulas, where $n = 2^m + s$ and $0 < s \leq 2^m$. The *Pseudobalanced Conjunction* $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ of $A_1, \ldots, A_n$ is defined inductively by:

- If $n = 1$ then $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ is just A_1.
- Otherwise $\bigwedge\!\!\!\bigwedge_{i=1}^{n} A_i$ is $\left(\bigwedge\!\!\!\bigwedge_{i=1}^{2^m} A_i\right) \wedge \left(\bigwedge\!\!\!\bigwedge_{i=1}^{s} A_{2^s+i}\right)$ and the first conjunct is balanced and the second is pseudobalanced.

Pseudobalanced formulas were first introduced by Bonet for the study of the number of symbols in propositional proofs [1, 2].

Lemma 19 *The formula* $\left(\bigwedge_{i=1}^{k-1} A_i\right) \wedge A_k \supset \bigwedge_{i=1}^{k} A_i$, *where the conjunctions are associated in a pseudobalanced way, has a tree-like Frege proof of* $O(\log k)$ *lines.*

Proof By induction on the depth of the formula $\bigwedge_{i=1}^{k} A_i$. (The depth is equal to $\lceil \log_2 k \rceil$.)

Base Case: The depth is one. In a constant number of lines we obtain a tree-like proof of $A_1 \wedge A_2 \supset A_1 \wedge A_2$.

Induction Step: Assume that the lemma holds for depth s. Let now $\bigwedge_{i=1}^{k-1} A_i \wedge A_k \supset \bigwedge_{i=1}^{k} A_i$ be a formula such that $\bigwedge_{i=1}^{k} A_i$ has depth $s+1$. Then $k = 2^s + a + 1$ where $0 \leq a < 2^s$. We consider separately the cases $a = 0$ and $a > 0$.

If $k - 1 = 2^s$, then $\bigwedge_{i=1}^{k-1} A_i \wedge A_k \supset \bigwedge_{i=1}^{k} A_i$ has a tree-like proof in constant number of lines, say c_1, since $\bigwedge_{i=1}^{k-1} A_i \wedge A_k$ is the same formula as $\bigwedge_{i=1}^{k} A_i$.

If $k - 1 = 2^s + a$ where $0 < a < 2^s$, then

$$(\bigwedge_{i=1}^{2^s} A_i \wedge \bigwedge_{i=2^s+1}^{2^s+a} A_i) \wedge A_k \supset \bigwedge_{i=1}^{2^s} A_i \wedge (\bigwedge_{i=2^s+1}^{2^s+a} A_i \wedge A_k)$$

has a tree-like Frege proof in a constant number of lines, say c_2. By the induction hypothesis

$$\bigwedge_{i=2^s+1}^{2^s+a} A_i \wedge A_k \supset \bigwedge_{i=2^s+1}^{2^s+a+1} A_i$$

has a tree-like proof in say cs lines. So with say c_3 more lines, we obtain

$$\bigwedge_{i=1}^{2^s} A_i \wedge (\bigwedge_{i=2^s+1}^{2^s+a} A_i \wedge A_k) \supset \bigwedge_{i=1}^{2^s} A_i \wedge \bigwedge_{i=2^s+1}^{2^s+a+1} A_i$$

and

$$\bigwedge_{i=1}^{k-1} A_i \wedge A_k \supset \bigwedge_{i=1}^{k} A_i$$

in a tree-like way. The last formula can be proven in $cs + c_2 + c_3$ lines. The result follows taking $c \geq c_1, c_2 + c_3$. $\square$

Lemma 20 *The formula* $\bigwedge_{i=1}^{k} A_i \supset \left(\bigwedge_{i=1}^{k} A_i\right) \wedge A_j$ *for* j *such that* $1 \leq j \leq k$ *has a tree-like Frege proof of* $O(\log k)$ *lines. Again, the conjunctions are associated in a pseudobalanced way.*

Proof By induction on the depth of $\bigwedge_{i=1}^{k} A_i$. The proof is similar to the proof of Lemma 19, and is left to the reader.

Lemma 21 *The formula* $\bigwedge_{i=1}^{k} A_i \supset \left(\bigwedge_{i=1}^{k} A_i\right) \wedge (A_j \wedge A_l)$ *where* $1 \leq j, l \leq k$ *and conjunctions are pseudobalanced, has a tree-like Frege proof of* $O(\log k)$ *lines.*

Proof By Lemma 20 the formulas

$$\bigwedge_{i=1}^{k} A_i \supset \bigwedge_{i=1}^{k} A_i \wedge A_j$$

and

$$\bigwedge_{i=1}^{k} A_i \supset \bigwedge_{i=1}^{k} A_i \wedge A_l$$

have proofs of $O(\log k)$ lines. With constantly many more lines we get the wanted result in $O(\log k)$ lines. $\square$

Definition The *height* of a proof is defined to be the largest integer h such that there exists formulas $B_1, \ldots, B_h$ in the proof with each B_{i+1} inferred by an inference which has B_i as a hypothesis.

Theorem 22 *If* $\vdash_{\overline{n}} A$ *then there is a tree-like Frege proof of* A *of* $O(n \log n)$ *lines and of height* $O(\log n)$.

Proof Let the Frege proof of A consist of the following sequence of lines:

$$A_1 \quad A_2 \quad \cdots \quad A_n.$$

Let B_i be the pseudobalanced conjunction $\bigwedge\!\!\!\bigwedge_{j=1}^{i} A_j$ where, by convention, B_0 is an arbitrary tautology (say, an axiom). We first show that for all $i < n$, the formula $B_i \supset B_{i+1}$ has a tree-like Frege proof of size $O(\log n)$ and hence of height $O(\log n)$. This is proved in two cases depending on how A_{i+1} is inferred.

Case 1: A_{i+1} is an axiom. Since A_{i+1} is an axiom there is a tree-like proof of

$$B_i \supset (B_i \wedge A_{i+1})$$

in a constant number of lines. Also, Lemma 19 gives us a treelike proof of the formula

$$(B_i \wedge A_{i+1}) \supset B_{i+1}$$

in $O(\log i)$ many lines. From these, the formula $B_i \supset B_{i+1}$ follows tautologically in a constant number of lines (namely, prove the tautology that the first two formulas imply the third formula in a constant size tree-like proof and then use modus ponens twice).

Case 2: A_{i+1} is obtained by modus ponens from former lines, say A_k and A_t. By Lemma 21, we obtain

$$B_i \supset B_i \wedge (A_k \wedge A_t)$$

in tree-like proof of $O(\log i)$ lines which is also trivially of height $O(\log i)$. There is a constant size tree-like proof of the tautology

$$B_i \wedge (A_k \wedge A_t) \supset B_i \wedge A_{i+1}.$$

And by Lemma 19 there is again a treelike proof of size and height $O(\log i)$ of

$$(B_i \wedge A_{i+1}) \supset B_{i+1}.$$

From these three formulas, $B_i \supset B_{i+1}$ follows tautologically with a constant size, tree-like proof.

We can now describe the treelike Frege proof of A. It begins with proofs of each of $B_i \supset B_{i+1}$ (in parallel). Without loss of generality, n is power of two, say $n = 2^s$ (if necessary, proof length can be padded by including extra axioms). Then for ℓ equal to 1, then 2, etc. up to s, the formulas

$$B_{i2^\ell} \supset B_{(i+1)2^\ell}$$

are proved, for all $0 \le i < 2^{s-\ell}$. The displayed formula is proved by a constant size tree-like proof from the formulas $B_{(2i)2^{\ell-1}} \supset B_{(2i+1)2^{\ell-1}}$ and $B_{(2i+1)2^{\ell-1}} \supset B_{(2i+2)2^{\ell-1}}$.

This then yields a $O(\log n)$ height and $O(n \log n)$ size proof of $B_0 \supset B_n$. Finally, since $A = A_n$, Lemma 20 implies there is a treelike proof of $B_n \supset A$ of size and height $O(\log n)$. Now the axiom B_0 and two further modus ponens inferences yield the formula A. □

As discussed at the end of the introduction, it is open whether the $O(n \log n)$ simulation in Theorem 22 can be improved, even if there are no restrictions on the height of the Frege proof. However, we can suggest a family of formulas which have linear size non-tree-like Frege proofs but might require size $\Omega(n \log n)$ size tree-like Frege proofs. As a special case of Corollary 3, any formula of the form

$$\bigwedge_{i=1}^{n} A_i \supset \bigwedge_{i=1}^{n} A_{j_i}$$

has a Frege proof of $O(n)$ lines. Examination of the proof of Corollary 3 shows that this Frege proof is not tree-like. We conjecture that there is a constant c such that these formulas require $\ge cn \log n$ line tree-like Frege proofs infinitely often (see [2] for an analysis of number of symbols in proofs of these formulas). This conjecture would imply the optimality of the simulation of Theorem 22.

An immediate consequence of Theorems 22 and Corollaries 15 and 18 is that tree-like Frege proofs simulate both natural deduction and the sequent calculus with an increase in size of $O(n\alpha(n)\log(n))$.

It is very easy to show that natural deduction and sequent calculus tree-like linearly simulate the tree-like Frege calculus, and we leave the proofs to reader. Then together with Theorem 22, we obtain the following corollaries.

Corollary 23 *If* $\vdash_n A$*, then* $\vdash^{\mathrm{PKT}}_{O(n \log n)} \longrightarrow A$*.*

Corollary 24 *If* $\vdash_n A$*, then* A *has a natural deduction proof with* $O(n \log n)$ *steps.*

These corollaries are surprising since PKT and natural deduction proofs are tree-like but Frege proofs are not.

References

[1] M. L. Bonet, *The Lengths of Propositional Proofs and the Deduction Rule*, PhD thesis, U.C. Berkeley, 1991.

[2] ——, *Number of symbols in Frege proofs with and without the deduction rule*, in Arithmetic, Proof Theory and Computational Complexity, P. Clote and J. Krajíček, eds., Oxford University Press, 1993, pp. 61–95.

[3] M. L. Bonet and S. R. Buss, *On the deduction rule and the number of proof lines*, in Proceedings Sixth Annual IEEE Symposium on Logic in Computer Science, 1991, pp. 286–297.

[4] ——, *On the serial transitive closure problem*, SIAM Journal on Computing, 24 (1995), pp. 109–122.

[5] S. R. Buss, *Bounded Arithmetic*, Bibliopolis, 1986. Revision of 1985 Princeton University Ph.D. thesis.

[6] S. A. Cook, *Feasibly constructive proofs and the propositional calculus*, in Proceedings of the Seventh Annual ACM Symposium on Theory of Computing, 1975, pp. 83–97.

[7] S. A. Cook and R. A. Reckhow, *The relative efficiency of propositional proof systems*, Journal of Symbolic Logic, 44 (1979), pp. 36–50.

[8] S. C. Kleene, *Introduction to Metamathematics*, Wolters-Noordhoff and North-Holland, 1971.

[9] J. Krajíček, *Lower bounds to the size of constant-depth Frege proofs*, Journal of Symbolic Logic, 59 (1994), pp. 587–598. IS THIS RIGHT?, DOES IT EXIST?

[10] V. P. Orevkov, *Upper bound on the lengthening of proofs by cut elimination*, Journal of Soviet Mathematics, 34 (1986), pp. 1810–1819. Original Russian version in Zap. Nauchn. Sem. L.O.M.I. Steklov **137**(1984)87-98.

[11] ——, *Reconstruction of a proof from its scheme*, Soviet Mathematics Doklady, 35 (1987), pp. 326–329. Original Russian version in Dokl. Akad. Nauk. **293** (1987) 313-316.

[12] T. Pitassi, P. Beame, and R. Impagliazzo, *Exponential lower bounds for the pigeonhole principle*, Computational Complexity, 3 (1993), pp. 97–140.

[13] D. Prawitz, *Ideas and results in proof theory*, in Proceedings of the Second Scandinavian Logic Symposium, 1971, pp. 235–307.

[14] R. A. Reckhow, *On the Lengths of Proofs in the Propositional Calculus*, PhD thesis, Department of Computer Science, University of Toronto, 1976. Technical Report #87.

[15] R. Statman, *Complexity of derivations from quantifier-free Horn formulae, mechanical introduction of explicit definitions, and refinement of completeness theorems*, in Logic Colloquium '76, R. Gandy and M. Hyland, eds., Amsterdam, 1977, North-Holland, pp. 505–517.

[16] M. Szabo, *The Collected Papers of Gerhard Gentzen*, North-Holland, 1969.

[17] G. Takeuti, *Proof Theory*, North-Holland, Amsterdam, 2nd ed., 1987.