HAL open science

# Unbounded Proof-Length Speed-up in Deduction Modulo

Guillaume Burel

▶ To cite this version:

Guillaume Burel. Unbounded Proof-Length Speed-up in Deduction Modulo. 2007. inria-00138195v1

HAL Id: inria-00138195

https://inria.hal.science/inria-00138195v1

Submitted on 23 Mar 2007 (v1), last revised 3 Jul 2007 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Unbounded Proof-Length Speed-up in Deduction Modulo

Guillaume Burel
Université Henri Poincaré & LORIA*
guillaume.burel@ens-lyon.org

inria-00138195, version 1 - 23 Mar 2007

## Abstract

*In 1973, Parikh proved a theorem conjectured by Gödel 37 years before, which says that it is possible to find arithmetical formulæ that are provable in first order arithmetic, but whose shorter proof in second order arithmetic is arbitrarily smaller than any proof in first order. On the other hand, resolution for higher order logic can be simulated step by step in a first order narrowing and resolution method based on deduction modulo, whose paradigm is to separate deduction and computation to make proofs clearer and shorter.*

*We first prove that it is possible to find formulæ whose smaller proof in natural deduction modulo is unboundedly smaller than any proof in pure natural deduction. Then, we show that a proof in the $i+1$-th order arithmetic can be transformed into a proof of linear length in the i-th order arithmetic modulo some finite terminating and confluent rewrite system. This allows us to prove that the speed-up conjectured by Gödel does not come from the deduction part of the proofs, but can be expressed as computation, therefore justifying the use of deduction modulo as an efficient first order setting simulating higher order.*

## 1. Introduction

Even if two logical systems are shown to be expressively equivalent, i.e. they can prove exactly the same formulæ, they can lead to very different proofs, in particular in terms of length. For instance, it is shown that Frege systems have an exponential speed-up over resolution for propositional logic [4]. However in mechanized theorem proving, the length of proofs has an importance: first, computers have limited capacities, and this can lead to a difference between the practical expressiveness of theoretically equivalent systems. Even if computing power is always increasing, so that one is no longer afraid to use SAT-solvers within verification tools (mainly because worst cases do not often occur in practice), it is not conceivable to build an automated theorem prover that produces proofs of non-elementary length. Second, the length of a proof is one (among others) criterion for defining the quality of a proof. Indeed, a smaller proof is often more readable and, in the case for instance of software certification and proof engineering, more communicable and in many cases also more maintainable. In [8, 1], this notion of "good proofs" is translated into a proof ordering, which of course may correspond to the comparison of proof lengths.

Obtaining a speed-up can also have a theoretical interest, because, as remarked by Parikh in the introductory paragraph of [15], "the celebrated P=NP? question can itself be thought of as a speed-up question." (See [6].) All this explains the research for new formalisms whose deductive systems provide smaller proofs, such as for instance the calculus of structures [3] w.r.t. the sequent calculus [13] (see [16]).

In this paper, the length of a proof will correspond to its number of steps (sometimes called lines), whatever the actual size of the formulæ appearing in them is. Considering the minimal length of the proof, the definition of a speed-up is the following: given some function $h$ over natural numbers, a system has a speed-up for $h$ over another one, if there exists infinite set of formulæ provable in both of them, such that, if the length of the proofs in the first system is $l$ and the length in the second system is $k$, then $k > h(l)$.

In 1936, Gödel [14] conjectured that there exists such a speed-up for all recursive functions between $i$-th order and $i+1$-th order arithmetic, no matter the formal system actually used. In other words, he stated that for all recursive functions $h$, it is possible to find an infinite set of formulæ such that, for each of them, denoted by $\Phi$, if $k$ is the minimal number of steps in the proofs of $\Phi$ in the $i$-th order arithmetic ($k$ is assumed to exist, so that $\Phi$ is provable in it), and $l$ is the minimal number of steps in the proofs of $\Phi$ in the $i+1$-th order arithmetic, then $k > h(l)$.

This result was proved for first-order arithmetic by Parikh [20], who actually proved a stronger theorem: this proof-length speed-up exists in fact also for non-recursive functions. This was generalized to all orders by Krajíček,

*UMR 7503 CNRS-INPL-INRIA-Nancy2-UHP

and was proved for the true language of arithmetic by Buss [5] (the former results used an axiomatization of arithmetic using ternary predicates to represent addition and multiplication). The theorem proved by Buss is stated as follow:

**THEOREM 1 ([5, THEOREM 3]).**
*Let $i \geq 0$. Then there is an infinite family $\mathcal{F}$ of $\prod_1^0$-formulæ such that*

1. *for all $P \in \mathcal{F}$, $Z_i \vdash P$*
2. *there is a fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $Z_{i+1} \vdash_{k\ steps} P$*
3. *there is no fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $Z_i \vdash_{k\ steps} P$.*

$Z_i$ corresponds to the $i+1$-th order arithmetic (so $Z_0$ is in fact first order arithmetic), and $Z_i \vdash_{k\ \text{steps}} P$ means that $P$ can be proved in at most $k$ steps within a schematic system — i.e. a Hilbert-type (or Frege) system with a finite number of axiom schemata and inference rules— for $i-1$-th order arithmetic. (In fact, Buss proved this theorem also for weakly schematic systems, i.e. schematic systems in which every tautology can be used as an axiom, as well as generalizations of axioms, but we will not use this fact here.) Because this theorem is concerned in arithmetic, an intuitive notion of computation take place in the proofs. Indeed, as remarked by Poincaré, proving that $2+2=4$ using the definition of the addition is just a verification, and not a demonstration, so that in a proof occur in fact not only pure deduction but also computation. Therefore, the question arises whether this speed-up comes from the deductive or the computational part of the proofs, or both of them.

Deduction modulo [10] is a presentation of a given logic —and the formalisms associated with it— identifying what corresponds to computation. The computational part of a proof is put in a congruence between formulæ modulo whom the application of the deduction rules takes place. This leads to the sequent calculus modulo and the natural deduction modulo. The congruence is better represented a set of rewrite rules that can rewrite terms but also *atomic propositions*: indeed, one wants for instance to consider the definition of the addition or multiplication using rewrite rules over terms as part of the computation, but also the following rewrite rule:

$$x \times y = 0 \ \rightarrow\ x = 0 \vee y = 0$$

which rewrites an atomic proposition to a formula, so that the following simple proof of $t \times t = 0$ can be deduced from a proof $\pi$ of $t = 0$:

$$\vee\text{-i}\ \frac{\begin{array}{c}\pi\\ t=0\end{array}}{t \times t = 0}\ {\scriptstyle t\times t=0 \ \longrightarrow\ t=0\vee t=0}$$

Deduction modulo is logically equivalent to the considered logic, but proofs are often considered as simpler, because the computation is hidden, letting the deduction clearly appear. Proofs are also claimed to be smaller for the same reason. Nevertheless, this fact was never quantified. This paper answers this issue.

Moreover, it is possible, in deduction modulo, to build proofs of Higher Order Logic using a first order system [9]. Using this, a step of higher order resolution is completely simulated by a step of ENAR, the resolution and narrowing method based on deduction modulo. Therefore, it seems reasonable to think that deduction modulo is able to give the same proof-length speed-ups as the ones occurring between $i+1$-th and $i$-th order arithmetic. This paper therefore investigates how to relate proof-length speed-ups in arithmetic with the computational content of the proofs.

First, a formal system modulo some rewrite rules has been shown to be logically equivalent to the same system without modulo, but using assumptions in a theory that is said to be compatible with the set of rewrite rules (a more formal definition will be given in Section 3). For instance, it is equivalent to prove some formula $P$ in natural deduction modulo the rewrite rule $A \rightarrow A \vee B$ or to prove $P$ under the assumption $B \Rightarrow A$ in pure natural deduction.

Our first main result is the fact that natural deduction modulo some (finite) rewrite system has an unbounded speed-up over pure natural deduction using assumptions of a finite compatible theory. This idea is formalized in Theorem 5 of Section 5.1. As a corollary, we will find a infinite set of tautologies that proves that natural deduction modulo has an unbounded speed-up over pure natural deduction (even without assumptions).

This result is proven (see Section 5.1) using a rewrite system quite different from arithmetic. To get something more near to it, the idea is to prove that deduction modulo allows to simulate the $i+1$-th order in $i$-th order arithmetic, so that one can achieve the same speed-up between $Z_i$ and $Z_i$ modulo some $\mathcal{R}_i$ than between $Z_i$ and $Z_{i+1}$. In other words, there exists some rewrite system $\mathcal{R}_i$ such that natural deduction modulo $\mathcal{R}_i$ using assumptions in $Z_i$ has an unbounded speed-up over natural deduction using assumptions in $Z_i$. This is translated by Theorem 7 of Section 5.2. To prove this result, we will need some translations between some particular schematic system for $i$-th order arithmetic and natural deduction modulo (or not) some rewrite system.

In a converse way, one can prove that working modulo such a rewrite system at the order $i$ allows to bypass Buss' theorem, i.e. we will not be able to find an infinite family of formulæ such that each of them is provable in $Z_i$, there is a bound to the length of the proofs in natural deduction using assumptions from $Z_{i+1}$ but there is no bound in natural deduction modulo $\mathcal{R}_i$ using assumptions from $Z_i$. We can prove in fact a stronger theorem which says that a proof

in natural deduction using assumptions from $Z_{i+1}$ can be translated linearly into a proof in natural deduction modulo $\mathcal{R}_i$ using assumptions from $Z_i$ and a finite number of extra axioms. This corresponds to Theorem 8 and Corollary 9 of Section 5.3.

In the next section, we will recall the definition of a schematic system, and we will present such a system for $i$-th order arithmetic. The section 3 will define formally what deduction modulo, and in particular natural deduction modulo consists of. In Section 4 we will give the exact translations between a proof in the schematic system for $i$-th order arithmetic and a proof in natural deduction, modulo or not. An upper bound of the increase in the length of the proofs due to these translations will be given. Finally, in Section 5 the proofs of the three theorem motivated above will be given, and we will conclude about the interest of working within a first-order system modulo to simulate higher order.

## 2. A schematic system for $i$-th order arithmetic

### 2.1. Schematic systems

We recall here, using Buss' notations [5], what a schematic system consists in. It is essentially an Hilbert-type (or Frege) proof system, i.e. valid formulæ are derived from a finite number of axiom schemata using a finite number of inference rules. Theorem 1 is true on condition that proofs are performed using a schematic system.

First, we recall how to build first order formulæ, mainly to introduce the notations we will use. A (first-order) many-sorted signature consists in a set of function symbols and a set of predicates, all of them with their arity (and co-arity for function symbols). We denote by $\mathcal{T}(\Sigma, V)$ the set of *terms* built from a signature $\Sigma$ and a set of variables $V$. An *atomic proposition* is given by a predicate symbol $A$ of arity $[i_1, \ldots, i_n]$ and by $n$ terms $t_1, \ldots, t_n \in \mathcal{T}(\Sigma, V)$ with matching sorts. It is denoted $A(t_1, \ldots, t_n)$. *Formulæ* can be built using the following grammar[1]:

$$\mathcal{P} \quad \stackrel{!}{=} \quad \bot \mid A \mid \mathcal{P} \wedge \mathcal{P} \mid \mathcal{P} \vee \mathcal{P} \mid \mathcal{P} \Rightarrow \mathcal{P} \mid \forall x.\, \mathcal{P} \mid \exists x.\, \mathcal{P}$$

where $A$ ranges over atomic propositions and $x$ over variables. $P \Leftrightarrow Q$ will be used as a syntactic sugar for $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$, as well as $\neg P$ for $P \Rightarrow \bot$ and $\Gamma \Rightarrow Q$ for $P_1 \Rightarrow \ldots \Rightarrow P_n \Rightarrow Q$ if $\Gamma = P_1, \ldots, P_n$. Positions in a term or a formula, free variables and substitutions are defined as usual. The replacement of a variable $x$ by a term $t$ in a formula $P$ is denoted by $\{t/x\}P$.

Then, given a many-sorted signature of first order logic, we can consider infinite sets of *metavariables* $\alpha^i$ for each sort $i$ (which will be substituted by variables), of *term variables* $\tau^i$ for each sort $i$ (which will be substituted by terms) and *proposition variables* $A(x_1, \ldots, x_n)$ for each arity $[i_1, \ldots, i_n]$ (which will be substituted by formulæ).

Metaterms are built like terms, except that they can contain metavariables and term variables. Metaformulæ are built like formulæ, except that they can contain proposition variables (which play the same role as predicates) and metaterms.

A *schematic system* is a finite set of inference rules, where an inference rule is a triple of a finite set of metaformulæ (the *premises*), a metaformulæ (the *conclusion*), and a set of side conditions of the forms $\alpha^j$ *is not free in* $\Phi$ or $s$ *is freely substitutable for* $\alpha^j$ *in* $\Phi$ where $\Phi$ is a metaformula and $s$ a metaterm of sort $j$. It is denoted by

$$\frac{\Phi_1 \qquad \cdots \qquad \Phi_n}{\Psi}\,(R)$$

An inference with an empty set of premises will be called an axiom schema.

### 2.2. $i$-th order arithmetic

$i$-th order arithmetic ($Z_{i-1}$) is a many-sorted theory with the sorts $0, \ldots, i-1$ and the signature

$$\begin{array}{lll} 0 : 0 & + : [0;0] \to 0 & = : [0;0] \\ s : [0] \to 0 & \times : [0;0] \to 0 & \in^j : [j; j+1] \end{array}.$$

The schematic system we use here consists of the following inference rules:

**Axiom schemata of classical logic.** We take the one used by Gentzen [13, Chapter 5] to prove the equivalence of his formalisms with an Hilbert-type proof system:

$$A \Rightarrow A \tag{1}$$

$$A \Rightarrow B \Rightarrow A \tag{2}$$

$$(A \Rightarrow A \Rightarrow B) \Rightarrow A \Rightarrow B \tag{3}$$

$$(A \Rightarrow B \Rightarrow C) \Rightarrow B \Rightarrow A \Rightarrow C \tag{4}$$

$$(A \Rightarrow B) \Rightarrow (B \Rightarrow C) \Rightarrow A \Rightarrow C \tag{5}$$

$$(A \wedge B) \Rightarrow A \tag{6}$$

$$(A \wedge B) \Rightarrow B \tag{7}$$

$$(A \Rightarrow B) \Rightarrow (A \Rightarrow C) \Rightarrow A \Rightarrow (B \wedge C) \tag{8}$$

$$A \Rightarrow (A \vee B) \tag{9}$$

$$B \Rightarrow (A \vee B) \tag{10}$$

$$(A \Rightarrow C) \Rightarrow (B \Rightarrow C) \Rightarrow (A \vee B) \Rightarrow C \tag{11}$$

$$\begin{gathered}(\forall \alpha^j.\, A(\alpha^j)) \Rightarrow A(\tau^j) \\ \big(\tau^j \text{ is freely substitutable for } \alpha^j \text{ in } A(\alpha^j)\big)\end{gathered} \tag{12}$$

$$\begin{gathered}A(\tau^j) \Rightarrow \exists \alpha^j.\, A(\alpha^j) \\ \big(\tau^j \text{ is freely substitutable for } \alpha^j \text{ in } A(\alpha^j)\big)\end{gathered} \tag{13}$$

$$A \vee (A \Rightarrow \bot) \tag{14}$$

[1] $\stackrel{!}{=}$ is used for definitions.

**Inference rules of classical logic.** Again, we take the one used by Gentzen [13]:

$$\frac{A \qquad A \Rightarrow B}{B} \tag{15}$$

$$\frac{A \Rightarrow B(\beta^j)}{A \Rightarrow \forall \alpha^j.\, B(\alpha^j)} \; (\beta^j \text{ is not free in } A \Rightarrow \forall \alpha^j.\, B(\alpha^j)) \tag{16}$$

$$\frac{B(\beta^j) \Rightarrow A}{(\exists \alpha^j.\, B(\alpha^j)) \Rightarrow A} \; (\beta^j \text{ is not free in } (\exists \alpha^j.\, B(\alpha^j)) \Rightarrow A) \tag{17}$$

**Identity axiom schemata.** They define the particular relation $=$:

$$\forall \alpha^0.\, \alpha^0 = \alpha^0 \tag{18}$$
$$\forall \alpha^0 \beta^0.\, \alpha^0 = \beta^0 \Rightarrow s(\alpha^0) = s(\beta^0) \tag{19}$$
$$\forall \alpha^0 \beta^0 \gamma^0.\, \alpha^0 = \beta^0 \Rightarrow \alpha^0 + \gamma^0 = \beta^0 + \gamma^0 \tag{20}$$
$$\forall \alpha^0 \beta^0 \gamma^0.\, \alpha^0 = \beta^0 \Rightarrow \gamma^0 + \alpha^0 = \gamma^0 + \beta^0 \tag{21}$$
$$\forall \alpha^0 \beta^0 \gamma^0.\, \alpha^0 = \beta^0 \Rightarrow \alpha^0 \times \gamma^0 = \beta^0 \times \gamma^0 \tag{22}$$
$$\forall \alpha^0 \beta^0 \gamma^0.\, \alpha^0 = \beta^0 \Rightarrow \gamma^0 \times \alpha^0 = \gamma^0 \times \beta^0 \tag{23}$$
$$\forall \alpha^0 \beta^0.\, \alpha^0 = \beta^0 \Rightarrow A(\alpha^0) \Rightarrow A(\beta^0) \tag{24}$$

**Robinson's axioms.** They are the axioms defining the function symbols of arithmetic [18]:

$$\forall \alpha^0.\, \neg\, s(\alpha^0) = 0 \tag{25}$$
$$\forall \alpha^0 \beta^0.\, s(\alpha^0) = s(\beta^0) \Rightarrow \alpha^0 = \beta^0 \tag{26}$$
$$\forall \alpha^0.\, (\neg\, \alpha^0 = 0) \Rightarrow \exists \beta^0.\, \alpha^0 = s(\beta^0) \tag{27}$$
$$\forall \alpha^0.\, \alpha^0 + 0 = \alpha^0 \tag{28}$$
$$\forall \alpha^0 \beta^0.\, \alpha^0 + s(\beta^0) = s(\alpha^0 + \beta^0) \tag{29}$$
$$\forall \alpha^0.\, \alpha^0 \times 0 = 0 \tag{30}$$
$$\forall \alpha^0 \beta^0.\, \alpha^0 \times s(\beta^0) = \alpha^0 \times \beta^0 + \alpha^0 \tag{31}$$

**Induction and comprehension axiom schemata.**

$$A(0) \Rightarrow \big(\forall \beta^0.\, A(\beta^0) \Rightarrow A(s(\beta^0))\big) \Rightarrow \forall \alpha^0.\, A(\alpha^0) \tag{32}$$

For all $0 \leq j < i-1$,

$$\exists \alpha^{j+1}.\, \forall \beta^j.\, \beta^j \in^j \alpha^{j+1} \Leftrightarrow A(\beta^j) \quad (\alpha^{j+1} \text{ is not free in } A) \tag{33}$$

From this point on, we will denote by $Z_{i-1} \vdash^{\mathsf{S}}_{k} P$ the fact that there exists a proof of $P$ of length at most $k$ in this schematic system, i.e. $P$ can be derived using at most $k$ instances of these inference rules.

# 3. Deduction modulo

## 3.1. Rewriting formulæ

In deduction modulo, formulæ are considered modulo some congruence defined by some rules that rewrite not only terms but also formulæ.

A *term rewrite rule* is the pair of terms $l, r$ such that all free variables of $r$ appear in $l$. It is denoted $l \to r$. A *term rewrite system* is a set of term rewrite rules.

A term $s$ can be rewritten to a term $t$ by a term rewrite rule $l \to r$ if there exists some substitution $\sigma$ and some position $\mathfrak{p}$ in $s$ such that $\sigma l = s_{|\mathfrak{p}}$ and $t = s[\sigma r]_{\mathfrak{p}}$.

An atomic proposition $A(s_1, \ldots, s_i, \ldots, s_n)$ can be rewritten to the atomic proposition $A(s_1, \ldots, t_i, \ldots, s_n)$ by a term rewrite rule $l \to r$ if $s_i$ can be rewritten to $t_i$ by $l \to r$. This relation is extended by congruence to all formulæ.

A *proposition rewrite rule* is the pair of an atomic proposition $A$ and a formula $P$, such that all free variables of $P$ appear in $A$. It is denoted $A \to P$. A *proposition rewrite system* is a set of proposition rewrite rules.

A formula $Q$ can be rewritten to a formula $R$ by a proposition rewrite rule $A \to P$ if there exists some substitution $\sigma$ and some position $\mathfrak{p}$ in $Q$ such that $\sigma A = Q_{|\mathfrak{p}}$ and $R = Q[\sigma P]_{\mathfrak{p}}$. Semantically, this proposition rewrite relation must be seen as a logical equivalence between formulæ.

The fact that $P$ can be rewritten to $Q$ either by a term or by a proposition rewrite rule of a rewrite system $\mathcal{R}$ will be denoted by $A \underset{\mathcal{R}}{\longrightarrow} P$. The transitive closure of these relation will be denoted by $\overset{*}{\underset{\mathcal{R}}{\longrightarrow}}$, its reflexive transitive closure by $\overset{*}{\underset{\mathcal{R}}{\longleftrightarrow}}$.

## 3.2. Natural deduction modulo

Using some equivalence $\overset{*}{\underset{\mathcal{R}}{\longleftrightarrow}}$ defined by a term and proposition rewrite system $\mathcal{R}$, we can define natural deduction modulo as in [11]. Its inference rules are represented in Figure 1. They are the same as the one introduced by Gentzen [13], except that we work modulo the rewrite relation. Leaves of a proof that are not introduced by some inference rules (contrary to $A$ in $\Rightarrow$-i for instance) are the assumptions of the proof. Note that if we do not work modulo, $\Rightarrow$-e is exactly the same as (15).

The length of a proof is the number of inferences used in it. We will denote by $\mathcal{T} \vdash^{\mathsf{N}}_{k\,\mathcal{R}} P$ the fact that there exists a proof of $P$ of length at most $k$ using a finite subset of $\mathcal{T}$ ($\mathcal{T}$ can be infinite) as assumptions. In the case where $\mathcal{R} = \emptyset$, we are back to pure natural deduction, and we will use $\mathcal{T} \vdash^{\mathsf{N}}_{k} P$. Abusing notations, we will write $Z_i \vdash^{\mathsf{N}}_{k\,\mathcal{R}} P$ to say that there is a proof of $P$ of length at most $k$ using as assumptions a finite subset of instances of the axiom schemata (18) to (33).

$$\Rightarrow\text{-i}\ \frac{\begin{matrix}[A]\\ B\end{matrix}}{C}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \Rightarrow B \qquad \Rightarrow\text{-e}\ \frac{A \qquad C}{B}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \Rightarrow B$$

$$\wedge\text{-i}\ \frac{A \qquad B}{C}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \wedge B \qquad \wedge\text{-e}\ \frac{C}{A}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \wedge B \text{ or } C \xleftrightarrow[\mathcal{R}]{*} B \wedge A$$

$$\vee\text{-i}\ \frac{A}{C}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \vee B \text{ or } C \xleftrightarrow[\mathcal{R}]{*} B \vee A \qquad \vee\text{-e}\ \frac{C \qquad \begin{matrix}[A]\\ D\end{matrix} \qquad \begin{matrix}[B]\\ D\end{matrix}}{D}\ \text{if } C \xleftrightarrow[\mathcal{R}]{*} A \vee B$$

$$\forall\text{-i}\ \frac{\{y/x\}A}{B}\ \text{if } B \xleftrightarrow[\mathcal{R}]{*} \forall x.\ A \text{ and } y \text{ is not free in } A \text{ nor in the assumptions of the proof above} \qquad \forall\text{-e}\ \frac{A}{B}\ \text{if } A \xleftrightarrow[\mathcal{R}]{*} \forall x.\ C \text{ and } B \xleftrightarrow[\mathcal{R}]{*} \{t/x\}C$$

$$\exists\text{-i}\ \frac{B}{A}\ \text{if } A \xleftrightarrow[\mathcal{R}]{*} \exists x.\ C \text{ and } B \xleftrightarrow[\mathcal{R}]{*} \{t/x\}C \qquad \exists\text{-e}\ \frac{B \qquad \begin{matrix}[\{y/x\}A]\\ C\end{matrix}}{C}\ \text{if } B \xleftrightarrow[\mathcal{R}]{*} \exists x.\ A \text{ and } y \text{ is not free in } C \text{ nor in the assumption of the proof above except } \{y/x\}A$$

$$\textit{classical}\ \frac{}{B}\ \text{if } A \xleftrightarrow[\mathcal{R}]{*} B \vee (B \Rightarrow \bot) \qquad \bot\text{-e}\ \frac{A}{B}\ \text{if } A \xleftrightarrow[\mathcal{R}]{*} \bot$$

**Figure 1. Inference Rules of Natural Deduction Modulo.**

Following Definition 1.4 of [10], a theory $\mathcal{T}$ is said *compatible* with a rewrite system $\mathcal{R}$ if:

- $P \xleftrightarrow[\mathcal{R}]{*} Q$ implies $\mathcal{T} \vdash^{\mathsf{N}} P \Leftrightarrow Q$;
- for every proposition $P \in \mathcal{T}$, we have $\vdash^{\mathsf{N}}_{\mathcal{R}} P$.

For instance, as stated in the introduction, $B \Rightarrow A$ is compatible with $A \rightarrow A \vee B$: it possible to prove $A \Leftrightarrow A \vee B$ assuming $B \Rightarrow A$ with the proof of Figure 2 (other cases of equivalent formulæ can be derived from it), and reciprocally, $B \Rightarrow A$ has the following proof modulo $A \rightarrow A \vee B$:

$$\Rightarrow\text{-i}\ \frac{\vee\text{-i}\ \dfrac{B\ \text{(i)}}{A}\ A \longrightarrow A \vee B}{B \Rightarrow A}\ \text{(i)}$$

Given a rewrite system, a compatible theory always exists, and one can show that proving modulo a rewrite system is the same as proving without modulo but using a compatible theory as assumptions [10, Proposition 1.8].

## 4. Translations

### 4.1. From $Z_i \vdash^{\mathsf{S}}$ to $Z_i \vdash^{\mathsf{N}}$

We want to translate a proof in the schematic system of $Z_i$ into a proof in pure natural deduction using as assumptions instances of the axiom schemata (18) to (33).

For the axiom schemata and inference rules of classical logic, we use the same translation as Gentzen, for instance the axiom schema (4) is translated into the natural deduction proof

$$\Rightarrow\text{-i}\ \frac{\Rightarrow\text{-i}\ \dfrac{\Rightarrow\text{-i}\ \dfrac{\Rightarrow\text{-e}\ \dfrac{B\ \text{(ii)} \qquad \Rightarrow\text{-e}\ \dfrac{A\ \text{(iii)} \qquad A \Rightarrow B \Rightarrow C\ \text{(i)}}{B \Rightarrow C}}{C}}{A \Rightarrow C}\ \text{(iii)}}{B \Rightarrow A \Rightarrow C}\ \text{(ii)}}{(A \Rightarrow B \Rightarrow C) \Rightarrow B \Rightarrow A \Rightarrow C}\ \text{(i)}$$

and the inference rule (17) into

$$\exists\text{-e}\ \frac{\exists \alpha^j.\ B(\alpha^j)\ \text{(i)} \qquad \Rightarrow\text{-e}\ \dfrac{B(\beta^j)\ \text{(ii)} \qquad B(\beta^j) \Rightarrow A}{A}}{\Rightarrow\text{-i}\ \dfrac{A}{\exists \alpha^j.\ B(\alpha^j) \Rightarrow A}\ \text{(i)}}\ \text{(ii)}$$

(note that the side condition ensure that it is possible to consider that what will be substituted for $\beta$ is free in $A$ and the assumptions of the proof above $B(\beta^j) \Rightarrow A$).

All these inference rules have a translation whose length does not depend on the formulæ finally substituted in the proof.

In a schematic system proof, there is also a finite number of instances of the axioms schemata for identity, Robinson's axioms and induction and comprehension schemata. We keep these instances as assumptions in natural deduction: let $E$ be the axioms of identity, except (24), $R$ Robinson's axioms, and $\Gamma_\pi$ the particular instances of (24), (32) and (33) used in $\pi$. Then, a proof $\pi$ of $P$ in the schematic system for $Z_i$ can be translated into a proof of $P$ in natural deduction using assumptions in $E$, $R$ and $\Gamma_\pi$, such that its length is linear compared to the length of $\pi$. We can therefore state the following proposition:

**Proposition 2.** *It is possible to translate a proof of length $n$ in the schematic system for $Z_i$ into a proof of length $O(n)$*

$$\wedge\text{-i}\;\dfrac{\Rightarrow\text{-i}\;\dfrac{\vee\text{-i}\;\dfrac{A\ (\text{i})}{A\vee B}}{A\Rightarrow A\vee B}\ (\text{i}) \qquad \Rightarrow\text{-i}\;\dfrac{\vee\text{-e}\;\dfrac{A\vee B\ (\text{ii}) \qquad A\ (\text{iii}) \qquad \Rightarrow\text{-e}\;\dfrac{B\ (\text{iii}) \qquad B\Rightarrow A}{A}}{A}\ (\text{iii})}{A\vee B\Rightarrow A}\ (\text{ii})}{A\Leftrightarrow A\vee B}$$

**Figure 2. Proof partly showing that $B\Rightarrow A$ is compatible with $A\rightarrow A\vee B$.**

*in (pure) natural deduction using assumptions in $Z_i$.*

$$Z_i \vdash^{\mathsf{S}}_{k} P \;\rightsquigarrow\; Z_i \vdash^{\mathsf{N}}_{O(k)} P$$

### 4.2. From $Z_i \vdash^{\mathsf{N}}$ to $Z_i \vdash^{\mathsf{S}}$

In these section, we consider a proof of $P$ in natural deduction, using as assumption $E$, $R$ and a finite set of instances of (24), (32) and (33) for $Z_i$. We translate it into a proof in the schematic system for $Z_i$.

This is essentially a generalization of the translation from the $\lambda$-calculus to combinatory logic (see [7]). We define mutually recursively two functions by induction on the inference rules: T transforms a proof of $P$ in natural deduction using assumptions $\Gamma$ into a proof of $P$ in the schematic system (1) to (17) plus $\Gamma$. $\mathrm{T}_A$ transform a proof of $P$ in natural deduction using assumptions $\Gamma, A$ into a proof of $A\Rightarrow P$ in the schematic system (1) to (17) plus $\Gamma$.

The definition of T and $\mathrm{T}_A$ are represented in Figure 3, although only in the cases for $\Rightarrow$ and $\exists$, by lack of place and because these are the most representative ones. The others can be guessed by the reader. The definition of $\mathrm{T}_A$ for $\Rightarrow$-i is not looping, because they are no longer $\Rightarrow$-i in $\mathrm{T}_B(\pi)$. Nevertheless, this case impose use to define what $\mathrm{T}_A$ means for a proof using the inference rules (16) and (17). (The translation of (15) is already defined because (15) is equal to $\Rightarrow$-e.)

It can be verified that this definition transforms a proof of size $n$ into a proof of size $O(3^n)$. Due to [6, Corollary 3.4], we could have found, at least for the propositional part, a polynomial translation. Nevertheless all we need in this paper is the fact that the increase of the proof length in the translation is bounded.

**Proposition 3.** *It is possible to translate a proof of length $n$ in the (pure) natural deduction using assumptions in $Z_i$ into a proof of length $O(3^n)$ in the schematic system for $Z_i$.*

$$Z_i \vdash^{\mathsf{N}}_{k} P \;\rightsquigarrow\; Z_i \vdash^{\mathsf{S}}_{O(3^k)} P$$

### 4.3. From $Z_{i+1} \vdash^{\mathsf{S}}$ to $Z_i \vdash^{\mathsf{N}}_{\mathcal{R}_i}$

This time, we translate a proof in the schematic system for $Z_{i+1}$ into a proof in natural deduction modulo using as assumption instances of the axiom schemata (18) to (33), but in the language of $Z_i$. The point is that using modulo it is possible to downshift an order.

We follow the translation of Section 4.1, except for the axiom schemata (24), (32) and (33) that are instantiated by formulæ that are in the language of $Z_{i+1}$ but not in the language of $Z_i$. To translate these schemata, we will use the work of F. Kirchner [17] which permits to express first-order theories using a finite number of axioms. The idea is to transform some metaformula $A(t_1,\ldots,t_n)$ used in an axiom schema into a formula of the form $\langle t_1,\ldots,t_n\rangle\ \epsilon\ \gamma$ where $\gamma$ will be some term representing what formula will be actually substituted for $A$.

Following F. Kirchner's method, we add new sorts $\ell$ for lists and $c$ for classes, as well as new function symbols and predicate

$$\begin{array}{rclrclrcl}
1^j & : & j & \dot{=} & : & [0;0]\rightarrow c & \emptyset & : & c \\
S^j & : & [j]\rightarrow j & \dot{\in}^j & : & [j;j+1]\rightarrow c & \mathcal{P}^j & : & [c]\rightarrow c \\
\cdot[\cdot]^j & : & [j;\ell]\rightarrow j & \cup & : & [c;c]\rightarrow c & \mathcal{C}^j & : & [c]\rightarrow c \\
nil & : & \ell & \cap & : & [c;c]\rightarrow c & \epsilon & : & [\ell;c] \\
::^j & : & [j;\ell]\rightarrow \ell & \supset & : & [c;c]\rightarrow c & & &
\end{array}.$$

$\langle\alpha_1,\ldots,\alpha_n\rangle$ will be syntactic sugar for $\alpha_1 ::^{j_1} \cdots :: \alpha_n ::^{j_n} nil$ for the appropriate $j_i$. We change the axiom schemata (24), (32) and (33) into the following *axioms*:

$$\forall\gamma^c.\ \forall\alpha^0\beta^0.\ \alpha^0=\beta^0 \Rightarrow \langle\alpha^0\rangle\ \epsilon\ \gamma^c \Rightarrow \langle\beta^0\rangle\ \epsilon\ \gamma^c \tag{34}$$

$$\forall\gamma^c.\langle 0\rangle\ \epsilon\ \gamma^c \Rightarrow \big(\forall\beta^0.\ \langle\beta^0\rangle\ \epsilon\ \gamma^c \Rightarrow \langle s(\beta^0)\rangle\ \epsilon\ \gamma^c\big) \Rightarrow \forall\alpha^0.\ \langle\alpha^0\rangle\ \epsilon\ \gamma^c \tag{35}$$

For all $0\le j<i-1$,

$$\forall\gamma^c.\ \exists\alpha^{j+1}.\ \forall\beta^j.\ \beta^j\in^j\alpha^{j+1} \Leftrightarrow \langle\beta^j\rangle\ \epsilon\ \gamma^c \tag{36}$$

The rewrite system $\mathcal{R}_i$ is then the following:

$$\mathrm{T}\left(\Rightarrow\text{-i}\,\dfrac{\pi\{\begin{smallmatrix}[A]\\ B\end{smallmatrix}}{A\Rightarrow B}\right) \stackrel{!}{=} \mathrm{T}_A\left(\pi\{\begin{smallmatrix}[A]\\ B\end{smallmatrix}\right)$$

$$\mathrm{T}\left(\Rightarrow\text{-e}\,\dfrac{\begin{smallmatrix}\pi_1\\ A\end{smallmatrix}\qquad\begin{smallmatrix}\pi_2\\ A\Rightarrow B\end{smallmatrix}}{B}\right) \stackrel{!}{=} (15)\,\dfrac{\begin{smallmatrix}\mathrm{T}(\pi_1)\\ A\end{smallmatrix}\qquad\begin{smallmatrix}\mathrm{T}(\pi_2)\\ A\Rightarrow B\end{smallmatrix}}{B}$$

$$\mathrm{T}\left(\exists\text{-i}\,\dfrac{\begin{smallmatrix}\pi\\ \{t/x\}A\end{smallmatrix}}{\exists x.\,A}\right) \stackrel{!}{=} (15)\,\dfrac{\begin{smallmatrix}\mathrm{T}(\pi)\\ \{t/x\}A\end{smallmatrix}\qquad\{t/x\}A\Rightarrow\exists x.\,A\ (13)}{\exists x.\,A}$$

$$\mathrm{T}\left(\exists\text{-e}\,\dfrac{\begin{smallmatrix}\pi_1\\ \exists x.\,A\end{smallmatrix}\qquad\pi_2\{\begin{smallmatrix}[A]\\ B\end{smallmatrix}}{B}\right) \stackrel{!}{=} (15)\,\dfrac{\begin{smallmatrix}\mathrm{T}(\pi_1)\\ \exists x.\,A\end{smallmatrix}\qquad(17)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi_2)\\ A\Rightarrow B\end{smallmatrix}}{(\exists x.\,A)\Rightarrow B}}{B}$$

$$\mathrm{T}(\mathrm{A}) \stackrel{!}{=} \mathrm{A}$$

$$\mathrm{T}_A\left(\Rightarrow\text{-i}\,\dfrac{\pi\{\begin{smallmatrix}[B]\\ C\end{smallmatrix}}{B\Rightarrow C}\right) \stackrel{!}{=} \mathrm{T}_A\left(\begin{smallmatrix}\mathrm{T}_B(\pi)\\ B\Rightarrow C\end{smallmatrix}\right)$$

$$\mathrm{T}_A\left(\Rightarrow\text{-e}\,\dfrac{\pi_1\{\begin{smallmatrix}[A]\\ B\end{smallmatrix}\qquad\pi_2\{\begin{smallmatrix}[A]\\ B\Rightarrow C\end{smallmatrix}}{C}\right) \stackrel{!}{=} (15)\,\dfrac{(15)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi_2)\\ A\Rightarrow B\Rightarrow C\end{smallmatrix}\qquad\cdots\ (4)}{B\Rightarrow A\Rightarrow C}\qquad(15)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi_1)\\ A\Rightarrow B\end{smallmatrix}\qquad\cdots\ (5)}{(B\Rightarrow A\Rightarrow C)\Rightarrow A\Rightarrow A\Rightarrow C}}{(15)\,\dfrac{A\Rightarrow A\Rightarrow C\qquad\cdots\ (3)}{A\Rightarrow C}}$$

$$\mathrm{T}_A\left(\exists\text{-i}\,\dfrac{\pi\{\begin{smallmatrix}[A]\\ \{t/x\}B\end{smallmatrix}}{\exists x.\,B}\right) \stackrel{!}{=} (15)\,\dfrac{\{t/x\}B\Rightarrow\exists x.\,B\ (13)\qquad(15)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi)\\ A\Rightarrow\{t/x\}B\end{smallmatrix}\qquad\cdots\ (5)}{(\{t/x\}B\Rightarrow\exists x.\,B)\Rightarrow A\Rightarrow\exists x.\,B}}{A\Rightarrow\exists x.\,B}$$

$$\mathrm{T}_A\left(\exists\text{-e}\,\dfrac{\pi_1\{\begin{smallmatrix}[A]\\ \exists x.\,B\end{smallmatrix}\qquad\pi_2\{\begin{smallmatrix}[A,B]\\ C\end{smallmatrix}}{C}\right) \stackrel{!}{=} (15)\,\dfrac{(17)\,\dfrac{\mathrm{T}_B\left(\begin{smallmatrix}\mathrm{T}_A(\pi_2)\\ A\Rightarrow C\end{smallmatrix}\right)}{\exists x.\,B\Rightarrow A\Rightarrow C}\qquad(15)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi_1)\\ A\Rightarrow\exists x.\,B\end{smallmatrix}\qquad\cdots\ (5)}{(\exists x.\,B\Rightarrow A\Rightarrow C)\Rightarrow A\Rightarrow A\Rightarrow C}}{(15)\,\dfrac{A\Rightarrow A\Rightarrow C\qquad\cdots\ (3)}{A\Rightarrow C}}$$

$$\mathrm{T}_A\left((17)\,\dfrac{\pi\{\begin{smallmatrix}[A]\\ B\Rightarrow C\end{smallmatrix}}{\exists x.\,B\Rightarrow C}\right) \stackrel{!}{=} (15)\,\dfrac{\begin{smallmatrix}\mathrm{T}_A(\pi)\\ A\Rightarrow B\Rightarrow C\end{smallmatrix}\qquad(A\Rightarrow B\Rightarrow C)\Rightarrow B\Rightarrow A\Rightarrow C\ (4)}{(15)\,\dfrac{(17)\,\dfrac{B\Rightarrow A\Rightarrow C}{\exists x.\,B\Rightarrow A\Rightarrow C}\qquad\cdots\ (4)}{A\Rightarrow\exists x.\,B\Rightarrow C}}$$

$$\mathrm{T}_A(A) \stackrel{!}{=} A\Rightarrow A\ (1)$$

$$\mathrm{T}_A\left(\begin{smallmatrix}\pi\\ B\end{smallmatrix}\right) \stackrel{!}{=} (15)\,\dfrac{\begin{smallmatrix}\mathrm{T}(\pi)\\ B\end{smallmatrix}\qquad B\Rightarrow A\Rightarrow B\ (2)}{A\Rightarrow B}$$

if the assumption $A$ is not actually used in $\pi$.

**Figure 3. Definition of the translation from $Z_i\vdash^{\mathsf{N}}$ to $Z_i\vdash^{\mathsf{S}}$.**

$$
\begin{array}{rcl}
t[nil]^j & \to & t \\
1^j[t ::^j l]^j & \to & t \\
S^j(n)[t ::^j l]^j & \to & n[l]^j \\
s(n)[l]^0 & \to & s(n[l]^0) \\
(t_1+t_2)[l]^0 & \to & t_1[l]^0 + t_2[l]^0 \\
(t_1 \times t_2)[l]^0 & \to & t_1[l]^0 \times t_2[l]^0 \\
l \in \doteq(t_1,t_2) & \to & t_1[l]^0 = t_2[l]^0 \\
l \in \dot{\in}^j(t_1,t_2) & \to & t_1[l]^j \in^j t_2[l]^{j+1} \\
l \in A \cup B & \to & l \in A \vee l \in B \\
l \in A \cap B & \to & l \in A \wedge l \in B \\
l \in A \supset B & \to & l \in A \Rightarrow l \in B \\
l \in \emptyset & \to & \bot \\
l \in \mathcal{P}^j(A) & \to & \exists x.\, x ::^j l \in A \\
l \in \mathcal{C}^j(A) & \to & \forall x.\, x ::^j l \in A
\end{array}
$$

Note that this system is convergent, i.e. terminating (either the size of a list decreases, or a $\cdot[\cdot]$ or an $\in$ goes more inside or disappears) and confluent (the only critical pairs, of the form:
$f(t_1,\ldots,t_n) \underset{\mathcal{R}_i}{\longleftarrow} f(t_1,\ldots,t_n)[nil] \underset{\mathcal{R}_i}{\longrightarrow} f(t_1[nil],\ldots,t_n[nil])$, are easily joinable).

Proposition 2 of [17] says that it is possible, for any formula $P$ of the language of $i$-th order arithmetic, to prove

$$\exists E.\, \forall x_1 \cdots x_n.\, \langle x_1,\ldots,x_n\rangle \in E \Leftrightarrow P \quad .$$

Moreover, the proof of this proposition show us how to construct the witness $E$. We will denote it by $E_P^{x_1,\ldots,x_n}$. Then, one can prove that $\langle t_1,\ldots,t_n\rangle \in E_P^{x_1,\ldots,x_n} \xrightarrow{*} \{t_1/x_1,\ldots,t_n/x_n\}P$. For instance, consider the formula $P \stackrel{!}{=} x = 0 \vee \exists y.\, x \in^0 y$. Then $E_P^x$ equals $\doteq(1,S(0)) \cup \mathcal{P}^1\left(\dot{\in}^0(S(1),1)\right)$ and $\langle t\rangle \in E_P^x$ can be rewritten to $t = 0 \vee \exists x.\, t \in^0 x$.

Consequently, the axiom schemata (24), (32) and (33) for formulæ of the language of $Z_{i+1}$ but not in the language of $Z_i$ are replaced by the proofs in Figure 4. In these translations, we need to instantiate $\gamma$ with some $E_A^x$. It is well-known that the instantiations are the most problematic rules in deductive systems, at least for automated provers. Nevertheless, the instantiation here is entirely and automatically determined by the formula used in the schema, so that no harm is done.

Using this, a proof $\pi$ of $P$ in the schematic system for $Z_{i+1}$ can be translated into a proof of $P$ in natural deduction modulo $\mathcal{R}_i$ using assumptions in $E$, $R$, (34), (35), (36) and $\Gamma'_\pi$ whose length is linear compared to the length of $\pi$, where $\Gamma'_\pi$ are the particular instances for the language of $Z_i$ of (24), (32) and (33) used in $\pi$.

**Proposition 4.** *It is possible to translate a proof of length $n$ in the schematic system for $Z_{i+1}$ into a proof of length $O(n)$ in the natural deduction modulo $\mathcal{R}_i$ using assumptions in $Z_i$, (34), (35) and (36).*

$$Z_{i+1} \vdash^{\mathsf{S}}_{k} P \;\rightsquigarrow\; Z_i, (34), (35), (36) \vdash^{\mathsf{N}}_{O(k)\ \mathcal{R}_i} P$$

## 5. Proof-length speed-ups

### 5.1. Speed-up over compatible theories

The following proves the existence of an unbounded speed-up (as in Theorem 1) for natural deduction modulo over pure natural deduction:

**THEOREM 5.**
*There is a rewrite system $\mathcal{R}$, there is an infinite family $\mathcal{F}$ such that such that for all finite compatible theories $\mathcal{T}$,*

1. *for all $P \in \mathcal{F}$, $\mathcal{T} \vdash^{\mathsf{N}} P$*
2. *there is a fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $\vdash^{\mathsf{N}}_{k\ steps\ \mathcal{R}} P$*
3. *there is no fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $\mathcal{T} \vdash^{\mathsf{N}}_{k\ steps} P$*

*Proof.* Consider the rewrite system $\mathcal{R}$:

$$s(x) + y \to x + s(y)$$

If $\underline{n}$ denotes the usual representation of the natural number $n$ using 0 and $s$, then it is quite clear that $\vdash^{\mathsf{N}}_{1\ \mathcal{R}} \underline{n} + \underline{n} = \underline{n+n}$. Let $\mathcal{T}$ be a finite theory compatible with $\mathcal{R}$. By definition $\mathcal{T} \vdash^{\mathsf{N}} \underline{n} + \underline{n} = \underline{n+n}$, but it is impossible to find a proof that takes less than $O(n)$ steps. (In the theory we may have some formulæ such as $s^m(x) + y = x + s^m(y)$ but they will only divides the minimal number of steps by $m$, and we can only have a finite number of such formulæ. The theorem is of course wrong if infinite theories are allowed, because one could add $\mathcal{F}$ to some theory compatible with $\mathcal{R}$ to get proofs with a bounded number of steps.) □

*Note:* We could also have used the system $\mathcal{R}_i$ of Section 4.3 and the formulæ of the form $\langle t_1,\ldots,t_n\rangle \in E_A^{x_1,\ldots,x_n} \Leftrightarrow \{t_1/x_1,\ldots,t_n/x_n\}A$.

**Corollary 6.** *There is a rewrite system $\mathcal{R}$ and an infinite family $\mathcal{F}$ such that*

1. *for all $P \in \mathcal{F}$, $\vdash^{\mathsf{N}} P$*
2. *there is a fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $\vdash^{\mathsf{N}}_{k\ steps\ \mathcal{R}} P$*
3. *there is no fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $\vdash^{\mathsf{N}}_{k\ steps} P$*

*Proof.* Consider $\mathcal{F}' \stackrel{!}{=} \{\mathcal{T} \Rightarrow P : P \in \mathcal{F}\}$ with the rewrite system $\mathcal{R}$ of Theorem 5, some finite theory $\mathcal{T}$ compatible with it, and the family $\mathcal{F}$ obtained in Theorem 5. By contradiction, if there is a $k$ such that for all $P' \in \mathcal{F}'$, $\vdash^{\mathsf{N}}_{k} P'$, then using some $\Rightarrow$-e, for all $P \in \mathcal{F}$, $\mathcal{T} \vdash^{\mathsf{N}}_{O(k)} P$. □

$$\forall\text{-e}\ \frac{\forall\gamma^{\mathcal{C}}.\ \forall\alpha^0\beta^0.\ \alpha^0=\beta^0 \Rightarrow \langle\alpha^0\rangle\ \epsilon\ \gamma^{\mathcal{C}} \Rightarrow \langle\beta^0\rangle\ \epsilon\ \gamma^{\mathcal{C}}\ (34)}{\forall\alpha^0\beta^0.\ \alpha^0=\beta^0 \Rightarrow A(\alpha^0) \Rightarrow A(\beta^0)}\ \langle\alpha^0\rangle\ \epsilon\ E_A^x \Rightarrow \langle\beta^0\rangle\ \epsilon\ E_A^x \xrightarrow{*} A(\alpha^0) \Rightarrow A(\beta^0)$$

$$\forall\text{-e}\ \frac{\forall\gamma^{\mathcal{C}}.\langle 0\rangle\ \epsilon\ \gamma^{\mathcal{C}} \Rightarrow \big(\forall\beta^0.\ \langle\beta^0\rangle\ \epsilon\ \gamma^{\mathcal{C}} \Rightarrow \langle s(\beta^0)\rangle\ \epsilon\ \gamma^{\mathcal{C}}\big) \Rightarrow \forall\alpha^0.\ \langle\alpha^0\rangle\ \epsilon\ \gamma^{\mathcal{C}}\ (35)}{A(0) \Rightarrow \big(\forall\beta^0.\ A(\beta^0) \Rightarrow A(s(\beta^0))\big) \Rightarrow \forall\alpha^0.\ A(\alpha^0)}\ \text{for all } t,\ \langle t\rangle\ \epsilon\ E_A^x \xrightarrow{*} A(t)$$

$$\forall\text{-e}\ \frac{\forall\gamma^{\mathcal{C}}.\ \exists\alpha^{j+1}.\ \forall\beta^j.\ \beta^j \in^j \alpha^{j+1} \Leftrightarrow \langle\beta^j\rangle\ \epsilon\ \gamma^{\mathcal{C}}\ (36)}{\exists\alpha^{j+1}.\ \forall\beta^j.\ \beta^j \in^j \alpha^{j+1} \Leftrightarrow A(\beta^j)}\ \langle\beta^j\rangle\ \epsilon\ E_A^x \xrightarrow{*} A(\beta^j)$$

**Figure 4. Translations of the axiom schemata (24), (32) and (33).**

### 5.2. Speed-up in arithmetic

We want to show that it is possible to achieve the same speed-up as the one between $i$-th order and $i+1$-th order arithmetic just by working modulo some rewrite system in $i$-th order arithmetic:

**THEOREM 7.**
*For all $i \geq 0$, there is a rewrite system $\mathcal{R}_i$ such that there is an infinite family $\mathcal{F}$ such that*

1. *for all $P \in \mathcal{F}$, $Z_i \vdash^{\mathsf{N}} P$*
2. *there is a fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $Z_i \vdash^{\mathsf{N}}_{k\ steps\ \mathcal{R}_i} P$*
3. *there is no fixed $k \in \mathbb{N}$ such that for all $P \in \mathcal{F}$, $Z_i \vdash^{\mathsf{N}}_{k\ steps} P$*

*Proof.* The rewrite system $\mathcal{R}_i$ is the one defined in Section 4.3. Let $\mathcal{F}$ be the family of formulæ obtained by Theorem 1. Let $\mathcal{F}' \stackrel{!}{=} \{(34) \Rightarrow (35) \Rightarrow (36) \Rightarrow P : P \in \mathcal{F}\}$. Then:

1. For all $P' \in \mathcal{F}'$, $Z_i \vdash^{\mathsf{N}} P'$: $Z_i \vdash^{\mathsf{S}} P$, therefore using Proposition 2, $Z_i \vdash^{\mathsf{N}} P$ and, adding to this proof three times $\Rightarrow$-i, $Z_i \vdash^{\mathsf{N}} P'$.
2. There is a $k$ such that for all $P' \in \mathcal{F}'$, $Z_i \vdash^{\mathsf{N}}_{k\ \mathcal{R}_i} P'$: there exists some $k$ such that for all $P \in \mathcal{F}$, $Z_{i+1} \vdash^{\mathsf{S}}_k P$. Using Proposition 4, there exists some $K$ such that for all $P \in \mathcal{F}$, we have $Z_i, (34), (35), (36) \vdash^{\mathsf{S}}_{K\ \mathcal{R}_i} P$ and one can add three $\Rightarrow$-i to obtain a proof of $P'$.
3. There is no $k$ such that for all $P' \in \mathcal{F}'$, $Z_i \vdash^{\mathsf{N}}_k P'$: Suppose by contradiction that there is a $k$ such that for all $P' \in \mathcal{F}'$, $Z_i \vdash^{\mathsf{N}}_k P'$, then using three times $\Rightarrow$-e, $Z_i, (34), (35), (36) \vdash^{\mathsf{N}}_{k+3} P$. But (34), (35) and (36) use function symbols not appearing in $P$ and $Z_i$ (for instance $\epsilon$). Therefore they cannot be used in a proof of $P$ in $Z_i$, so that in fact $Z_i \vdash^{\mathsf{N}}_{k+3} P$. Then, using Proposition 3, $Z_i \vdash^{\mathsf{S}}_{O(3^k)} P$, and that will be in contradiction with the fact that there is no $K$ such that for all $P$, $Z_i \vdash^{\mathsf{S}}_K P$.

If $\Gamma = (34), (35), (36)$,

$$\begin{array}{ccccc}
Z_{i+1} \vdash^{\mathsf{S}}_k P & \overset{\text{Prop. 4}}{\rightsquigarrow} & Z_i, \Gamma \vdash^{\mathsf{N}}_{K\ \mathcal{R}_i} P & \rightsquigarrow & Z_i \vdash^{\mathsf{N}}_{K+3\ \mathcal{R}_i} P' \\
\text{Theo. 1} \updownarrow & & & & \\
Z_i \vdash^{\mathsf{S}}_{3^k} P & \underset{\text{Prop. 3}}{\overset{\text{Prop. 2}}{\rightleftarrows}} & Z_i, \Gamma \vdash^{\mathsf{N}}_{k'} P & \rightsquigarrow & Z_i \vdash^{\mathsf{N}}_{k'} P' \quad \square
\end{array}$$

### 5.3. Bypassing Buss' speed-up using modulo

The goal of these section is to prove that one can work in $Z_i$ modulo some rewrite system $\mathcal{R}_i$ to be able to build proof as small as the one of $Z_{i+1}$. In fact, one can prove a stronger theorem:

**THEOREM 8.**
*For all $i \geq 0$, there exists a (finite) rewrite system $\mathcal{R}_i$ and a finite set of axioms $\Gamma$ such that for all formulæ $P$, if $Z_{i+1} \vdash^{\mathsf{N}}_k P$ then $Z_i, \Gamma \vdash^{\mathsf{N}}_{O(k)\ \mathcal{R}_i} P$.*

*Proof.* Let $\mathcal{R}_i$ be the rewrite system of Section 4.3 and $\Gamma \stackrel{!}{=} \{(34), (35), (36)\}$. We use the same idea as in Section 4.3, and we replace the instance of the axiom schemata (24), (32) and (33) by the axioms (34), (35) and(36) as indicated in Figure 4. $\square$

This results permits to make Gödel's theorem wrong if one works modulo $\mathcal{R}_i$.

**Corollary 9.** *For all $i \geq 0$, there exists a (finite) rewrite system $\mathcal{R}_i$ and a finite set of axioms $\Gamma$ such that there is no infinite family $\mathcal{F}$ of $\prod^0_1$-formulæ such that*

1. *for all $P \in \mathcal{F}$, $Z_i \vdash^{\mathsf{N}} P$*

2. *there is a fixed* $k \in \mathbb{N}$ *such that for all* $P \in \mathcal{F}$, $Z_{i+1} \vdash^{\mathsf{N}}_{k\ steps} P$
3. *there is no fixed* $k \in \mathbb{N}$ *such that for all* $P \in \mathcal{F}$, $Z_i, \Gamma \vdash^{\mathsf{N}}_{k\ steps\ \mathcal{R}_i} P$

The fact to add the finite set of axioms $\Gamma$ could be seen as some deceit, because we do not work really in $Z_i$, but in a theory strictly stronger. (By the way, due to Theorem 8, it is possible to prove the consistency of $Z_i$ in $Z_i, \Gamma$ modulo $\mathcal{R}_i$.) Nevertheless, the point here is that it is possible, by working modulo $\mathcal{R}_i$, to simulate $Z_{i+1}$ using a finite set of axioms, and not axiom schemata, without exploding the length of the proofs. If we were not working modulo, then it would not be possible to give a bound to the translation, because as noted after Theorem 5, to translate an axiom schema, we will have to introduce the adequate $E_A^{x_1,\ldots,x_n}$ and then, using a theory compatible with $\mathcal{R}_i$, decompose each step of rewriting. Therefore, the length of the translation will depend on the depth of the formulæ substituted in the axiom schemata.

It could also have been possible to translate the formulæ that one wants to prove, as is done in [12], where a formula of first order arithmetic is transformed by adding the information that some variable $n$ is an integer using some predicate $N(n)$ which can be rewritten into an axiom corresponding to the induction schema for first order arithmetic. Here, $P$ could be translated into $(34) \Rightarrow (35) \Rightarrow (36) \Rightarrow P$.

## 6. Conclusion and perspectives

We have first proved that, even with a very simple rewrite system, one can obtain in deduction modulo proofs of some tautologies that are unboundedly shorter than the proofs without modulo. This shows the power of separating computation and deduction. Of course, to actually find the proofs in deduction modulo, one will need to perform the computation, but the point is that this is more automatic and easier than the deduction itself.

Our second result is that it is possible to use some rewrite system to simulate the difference between $i$-th and $i+1$-th order arithmetic. This simulation allows to get the same proof speed-up for deduction modulo over non modulo systems than the one proved by Parikh.

Finally, we also proved that this simulation can be linear in term of proof length, at the condition to add three extra axioms which replace the missing axiom schemata. Together with the second result, this proves that the gap between $i$-th and $i+1$-th order arithmetic is in fact due to the computational part of the proofs. In this particular case, we also clearly identify the computation occurring in the proofs with a finite, convergent (so, in a sense, deterministic) rewrite system. This is not surprising, because, if one looks carefully, the proof of Theorem 1 given by Buss in [5] deeply relies on the fact that it is possible to define some truth predicate for the formulæ of the preceding order. Therefore, in a sense, it is possible, in $i+1$-th order arithmetic, to compute the validity of a formula in $i$-th order arithmetic.

These results are encouraging indicators that it is as good to work directly in higher order logics, as is done in the current interactive theorem provers, such as Coq [21] or Isabelle/HOL [19], or using a first order implementation of these logics, as could be done in a proof assistant based on deduction modulo (or on its sequel named superdeduction as in [2]). This paper gives clues to answer positively this question, although we were interested in the step between $i$-th order and $i+1$-th order arithmetic, and not between first order and higher order logic. The fact that higher order resolution can be simulated step by step by ENAR [9] is not a solution, because there may exist some other higher order proof system that produce proofs that cannot be conveniently translated in a first order system modulo. So, our next challenge will be, starting from the current results, to investigate how exactly higher order logic prevails or not over first order logic, by studying more closely the simulation of higher order logic.

A first direction to do so will be to prove that it is possible to apply transitivity between the simulation of $Z_{i+1}$ in $Z_i$ and the one of $Z_{i+2}$ in $Z_{i+1}$, in order to get a simulation of $Z_{i+2}$ in $Z_i$, for instance by combining $\mathcal{R}_i$ and $\mathcal{R}_{i+1}$. In addition to the expression of first order arithmetic as a theory modulo [12], this would lead to the expression of higher order arithmetic entirely as a theory modulo.

**Acknowledgments.** The author wishes to thank C. Kirchner and T. Hardin for many discussions and comments about this paper.

## References

[1] M. P. Bonacina and N. Dershowitz. Abstract canonical inference. *ACM Trans. Comput. Logic*, 8(1), 2007.

[2] P. Brauner. Un calcul des séquents extensible. Master's thesis, Université Henri Poincaré – Nancy 1, 2006. Website at: `http://rho.loria.fr/lemuridae.html`.

[3] K. Brünnler. *Deep Inference and Symmetry in Classical Proofs*. PhD thesis, Technische Universität Dresden, 2003.

[4] S. R. Buss. Polynomial size proofs of the propositional pigeonhole principle. *The Journal of Symbolic Logic*, 52(4):916–927, 1987.

[5] S. R. Buss. On Gödel's theorems on lengths of proofs I: Number of lines and speedups for arithmetic. *The Journal of Symbolic Logic*, 39(3):737–756, 1994.

[6] S. A. Cook and R. A. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44(1):36–50, 1979.

[7] H. B. Curry, R. Feys, and W. Craig. *Combinatory Logic*, volume 1. Elsevier Science Publishers B. V. (North-Holland), Amsterdam, 1958.

[8] N. Dershowitz and C. Kirchner. Abstract Canonical Presentations. *Theoretical Computer Science*, 357:53–69, 2006.

[9] G. Dowek, T. Hardin, and C. Kirchner. HOL-$\lambda\sigma$ an intentional first-order expression of higher-order logic. *Mathematical Structures in Computer Science*, 11(1):1–25, 2001.

[10] G. Dowek, T. Hardin, and C. Kirchner. Theorem proving modulo. *Journal of Automated Reasoning*, 31(1):33–72, 2003.

[11] G. Dowek and B. Werner. Proof normalization modulo. *The Journal of Symbolic Logic*, 68(4):1289–1316, 2003.

[12] G. Dowek and B. Werner. Arithmetic as a theory modulo. In J. Giesl, editor, *RTA*, volume 3467 of *Lecture Notes in Computer Science*, pages 423–437. Springer-Verlag, 2005.

[13] G. Gentzen. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1934. Translated in Szabo, editor., *The Collected Papers of Gerhard Gentzen* as "Investigations into Logical Deduction".

[14] K. Gödel. Über die Länge von Beweisen. *Ergebnisse eines Mathematischen Kolloquiums*, 7:23–24, 1936. English translation in [15].

[15] K. Gödel. On the length of proofs. In S. Feferman et al., editors, *Kurt Gödel: Collected Works*, volume 1, pages 396–399. Oxford University Press, Oxford, 1986.

[16] A. Guglielmi. Polynomial size deep-inference proofs instead of exponential size shallow-inference proofs. Available at `http://cs.bath.ac.uk/ag/p/AG12.pdf`, 2004.

[17] F. Kirchner. A finite first-order theory of classes. Available at `http://www.lix.polytechnique.fr/Labo/Florent.Kirchner/doc/fotc2006.pdf`, 2006.

[18] A. Mostowski, R. M. Robinson, and A. Tarski. *Undecidable Theories*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1953.

[19] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer, 2002.

[20] R. J. Parikh. Some results on the length of proofs. *Transactions of theACM*, 177:29–36, 1973.

[21] The Coq Development Team. *The Coq Proof Assistant Reference Manual*. INRIA, 2006. Version 8.0, available at `http://coq.inria.fr/doc/main.html`.