Annals of Mathematics and Artificial Intelligence 6 (1992) 107–126

On induction-free provability

Jan Krajíček [a,b] and Gaisi Takeuti [a]

[a] *University of Illinois at Urbana, Urbana, IL 61801, USA*
[b] *Mathematical Institute at Prague, Czechoslovakia*

0. Introduction

In this article we study relations between various fragments of bounded arithmetic. The fragments of interest here are the theories S_2^i and T_2^i introduced by Buss in [1]. The reader may recall that the principal axioms of S_2^i resp. of T_2^i are Σ_i^b-PIND resp. Σ_i^b-IND axioms, where the former are instances of polynomially bounded inducton for Σ_i^b-formulas while the latter are instances of ordinary induction for Σ_i^b-formulas.

Indications of the significance of these theories for complexity theory are two results due to Buss [1] and Krajíček et al. [7]:

(1) In [1] it is shown that a function is computable in polynomial time by an oracle machine quering a Σ_{i-1}^p-oracle iff it is Σ_i^b-definable in S_2^i. In particular, polynomial-time functions are precisely those Σ_1^b-definable in S_2^1.

(2) In [7] it is shown that $T_2^i = S_2^{i+1}$ implies that the polynomial hierarchy collapses to the $i+2$ level.

These statements show that it is an important problem to establish a relation between the fragments. It is easy to show that $S_2^i \subseteq T_2^i \subseteq S_2^{i+1}$ for $i \geqslant 1$. In [2] it is proved that S_2^{i+1} is $\forall\Sigma_{i+1}^b$-conservative over T_2^i, and by (2) above (as S_2^{i+1} is $\forall\Sigma_{i+2}^b$-axiomatized) it is not $\forall\Sigma_{i+2}^b$-conservative if $\Sigma_{i+2}^p \neq \Pi_{i+2}^p$.

In this paper we are interested in the relation of T_2^i to S_2^i, whether they are equal or different theories and, assuming that they are different, whether T_2^i is at least somehow conservative over S_2^i. This is a natural suspicion one gets after several unsuccessful attempts to separate these fragments of bounded arithmetic. It is possible that $S_2^i \neq T_2^i$ and still that T_2^i is $\forall\Sigma_i^b$-conservative over S_2^i. Note that this seems unlikely as it implies, in particular, that all functions Σ_1^b-definable in S_2 are polynomial time computable, a rather powerful statement with strong consequences for complexity theory, cf. [2].

The relation of S_2^i and T_2^i was also studied in [6]. It was shown there that the set of $\forall\Pi_1^b$-consequences of T_2^i is axiomatized over S_2^1 by a single $\forall\Pi_1^b$-sentence, namely $\mathrm{Con}(G_i)$ – consistency of propositional calculus G_i. Thus T_2^i is $\forall\Pi_1^b$-con-namely $\mathrm{Con}(G_i)$ – consistency of propositional calculus G_i. Thus T_2^i is $\forall\Pi_1^b$-conservative over S_2^i iff S_2^i proves $\mathrm{Con}(G_i)$.

If one tries to show that S_2^i does not prove $\mathrm{Con}(G_i)$ via some Gödel type argument the difficulty arises that the metanotion of provability (S_2^i-provability) is not the same as the notion referred to in the consistency statement (G_i-provability). Here we show that this obstacle can to a large extent be removed. Namely: provability in G_i is equivalent (in S_2^1) to a certain restricted provability in T_1^i which in turn is equivalent to a restricted induction-free provability in BASIC.

As a simple consequence we get – via known results – that T_{k+1}^i is not $\forall\Pi_1^b$-conservative over T_k^i ($i, k \geqslant 1$) and S_{k+1}^i is not $\forall\Pi_1^b$-conservative over S_k^i ($i, k \geqslant 2$).

In [6] the problem of conservativity of T_2^i over S_2^i was equivalently restated as a polynomial simulation problem of G_{i-1} versus G_i. Here we show that this question is also equivalent to a problem about lengthening of G_i-proofs after putting them into a tree-form. More specifically, consider the formula, $\mathrm{TREE}(G_i)$:

"$\forall d, A\ \exists d'$, if d is an G_i-proof of A then d' is a G_i-proof of A and d' is in a tree-form".

Then we have for $i \geqslant 1$ (proposition 4.1):

$$T_2^i \succcurlyeq_{\Sigma_i^b} S_2^i \quad \text{iff } S_2^1 \vdash \mathrm{TREE}(G_i).$$

(A proof is in a tree-form iff each sequent in it is used at most once as a hypothesis of an inference.)

Above we have used the notion of G_i-proof and other notions defined in [6]. Knowledge of these notions is not needed for a larger part of this paper and therefore we shall not repeat the definitions here. Otherwise we assume only basic knowledge of bounded arithmetic, see [1].

Some ideas from [4,5,10] are also used but no specific familiarity with these papers is assumed.

1. Preliminaries

The substitution rule:

$$\frac{\Gamma(b) \to \Delta(b)}{\Gamma(t) \to \Delta(t)},$$

allows to infer from a sequent its substitution instance obtained by substituting term t for all occurrences of free variable b. It is, of course, a derived rule of LKB. We shall often use it in the construction of LKB-proofs. However, we need to show that it can be eliminated from proofs without superpolynomial prolongation and without an increase of quantifier complexity, provided Γ or Δ contain a bounded existential quantifier (which will always be the case).

This is done as follows. Assume

$$\Gamma(b) \to \Delta(b) \tag{1}$$

is provable in LKB.

Derive:

$$b = t,\ \Gamma(t) \to \Delta(t),\ \Gamma(b) \tag{2}$$

and by cut with (1) get:

$$b = t,\ \Gamma(t) \to \Delta(t),\ \Delta(b). \tag{3}$$

Derive also:

$$b = t,\ \Delta(b),\ \Gamma(t) \to \Delta(t) \tag{4}$$

and by cut with (3) get:

$$b = t,\ \Gamma(t) \to \Delta(t). \tag{5}$$

From (5) follows

$$(\exists x \leqslant t,\ x = t),\ \Gamma(t) \to \Delta(t). \tag{6}$$

As also:

$$\to (\exists x \leqslant t,\ x = t) \tag{7}$$

is provable, cut with (6) gives required:

$$\Gamma(t) \to \Delta(t). \tag{8}$$

Note that proofs of (2), (4), (7) are easily constructed of length polynomial in the length of Γ, Δ, t, in a tree form and with the quantifier complexity that of Γ, Δ. Clearly the whole derivation can also be performed in S_2^1. Thus we can freely use the substitution rule as a rule of LKB, even when working in S_2^1.

To simplify proof-theoretic arguments we shall use the rather technical class of strictly Σ_i^b-formulas (to be defined below) instead of Σ_i^b-formulas. For this we need to fix a #-free, Σ_1^b-formula which is Σ_1^b-universal in S_2^1. Call such a fixed formula UNIV(a). Hence for every Σ_1^b-formula $B(\bar{x})$ we have $n < \omega$ such that $S_2^1 \vdash \forall \bar{x},\ B(\bar{x}) \equiv \mathrm{UNIV}(\langle \underline{n}, \bar{x} \rangle)$.

DEFINITION 1.1

For $i \geqslant 1$, a formula $A(a)$ is *strictly* Σ_i^b-, sΣ_i^b for short, iff it has the form:

$$\exists x_1 \leqslant t_1(a) \forall x_2 \leqslant t_2(a, x_1) \ldots Q_j x_j \leqslant t_j(a, x_1, \ldots, x_{j-1}) \ldots B(a, \bar{x}),$$

where $j = 1, \ldots, l < i$ and Q_j is $\exists$ iff j is odd, and $B(a, \bar{x})$ is UNIV($\langle a, \bar{x} \rangle$) if i is odd, respectively $\neg$UNIV($\langle a, \bar{x} \rangle$) if i is even. □

The following lemma is obvious.

LEMMA 1.2

Let X be the set of all #-free, $s\Sigma_i^b$-formulas provable in T_2^i ($i \geq 1$). Then $S_2^1 + X$ proves all $\forall\Sigma_i^b$-consequences of T_2^i. □

We shall need a certain form of elimination of the function symbol # from proofs. This is provided by the next lemma.

LEMMA 1.3

Let $A(a)$ be a #-free, $s\Sigma_i^b$-formula and assume:

$$T_2^i \vdash A(a).$$

Then for some $k < \omega$ the sequent:

$$2 \leqslant c_0, |a| \leqslant |c_0|, |c_0||c_0| \leqslant |c_1|, \ldots, |c_0||c_{k-1}| \leqslant |c_k \rightarrow A(a)$$

has #-free T_2^i-proof consisting of $s\Sigma_i^b$-formulas only and $c_0, \ldots, c_k$ are new free variables not occurring in $A(a)$.

Proof

The lemma follows by cut-elimination for T_2^i, compactness argument and the observation that for any term $t(a)$ there is $k < \omega$ such that

$$2 \leqslant c_0, |a| \leqslant |c_0|, |c_0||c_0| \leqslant |c_1|, \ldots, |c_0||c_{k-1}| \leqslant |c_k| \rightarrow t(a) \leqslant c_k$$

is provable in BASIC. □

The following probability notion was used in [4,5,10]. It defines a class of T_2^i-proofs with special properties.

DEFINITION 1.4

A triple $D = \langle d, \bar{t}, \bar{d}' \rangle$ is *i-regular proof* of sequent:

$$\Gamma(\bar{a}) \rightarrow \Delta(\bar{a})$$

iff the following conditions are satisfied:

(i) d is a sequence of sequents correctly inferred from previous ones using rules of LKB and Σ_i^b-IND and the end-sequent of d is $\Gamma(\bar{a}) \rightarrow \Delta(\bar{a})$;

(ii) every formula in d is #-free and a subformula of a $s\Sigma_i^b$-formula;

(iii) d is in a free variable normal form;

(iv) if $\bar{a}$ are all parameter variables in d and $\bar{b} = (b_0, \ldots, b_k)$ are all other free variables in d then: if the elimination rule of b_u is below the elimination rule of b_v then $u < v$, and the elimination rule of b_u ($u \leqslant k$) is either:

(a) Σ_i^b-IND:

$$\frac{A(b_u), \Gamma \rightarrow \Delta, A(s(b_u))}{A(0), \Gamma \rightarrow \Delta, A(r(b_0, \ldots, b_{u-1}, \bar{a}))}$$

or

(b) $\exists\leqslant$:left:

$$\frac{b_u \leqslant r(b_0, \quad, b_{u-1}, \bar{a}),\ A(b_u),\ \Gamma \to \Delta}{\exists x \leqslant r(b_0, \ldots, b_{u-1}, \bar{a}) A(x),\ \Gamma \to \Delta}$$

or

(c) $\forall\leqslant$:right:

$$\frac{b_u \leqslant r(b_0, . \quad, b_{u-1}, \bar{a}),\ \Gamma \to \Delta,\ A(b_u)}{\Gamma \to \Delta,\ \forall x \leqslant r(b_0, \ldots, b_{u-1}, \bar{a}) A(x)},$$

(v) $\bar{t}$ of D is a sequence of #-free terms containing terms $t_u(\bar{a})$ $(u \leqslant k)$; and it holds:

$$b_0 \leqslant t_0(\bar{a}), \ldots, b_{u-1} \leqslant t_{u-1}(\bar{a}) \to$$
$$\to r(b_0, \ldots, b_{u-1}, \bar{a}) \leqslant t_u(\bar{a}), \qquad (*)_u$$

where $r(b_0, \ldots, b_{u-1}, \bar{a})$ is the term from the elimination rule of b_u (cf. (iv));

(vi) $\bar{d}'$ is a sequence of proofs containing proofs $d'_u (u \leqslant k)$ of $(*)_u$ which are induction-free, quantifier-free and contains only variables $b_0, \ldots, b_{u-1}, \bar{a}$. These proofs are called supplementary. □

Note that we do not require d to be in a tree form, i.e. a sequent may be an upper sequent of more than one inference.

DEFINITION 1.5

(a) i-RPr(a, b) is a formalization of:
"there is i-regular proof $D \leqslant a$ of b".

(b) i-RPr*(a, b) is a formalization of:
"there is i-regular proof $D \leqslant a$ of b which is in a tree form".

(c) i-IFRPr(a, b) is a formalization of:
"there is i-regular, induction-free proof $D \leqslant a$ of b".

(d) i-IFRPr*(a, b) is a formalization of:
"there is i-regular, induction-free proof $D \leqslant a$ of b which is in a tree form".

These formalizations are taken to be Δ_1^b w.r.t. S_2^1. We also assume that all formulas defined above are subformulas of UNIV – this requirement is easy to meet (by possible enlarging of UNIV) – and it implies that all these formulas are $s\Sigma_1^b$. □

Recall also the definition of dyadic numerals.

DEFINITION 1.6

A dyadic numeral of n, denoted $\underline{n}$, is inductively defined:

$$\underline{0} := 0,\ \underline{1} := 1,\ \underline{2} := (1+1),\ \underline{2n} := (\underline{2} \cdot \underline{n}) \quad \text{and } \underline{2n+1} := s(\underline{2n}).$$

The formalization of the dyadic numeral in S_2^1 is denoted $\underset{\sim}{a}$. □

Finally we state a form of Σ_1^b-completeness, a well-known observation, cf. [1 chapter 7] or [9, chapter 6].

LEMMA 1.7

For any #-free, sΣ_1^b-formula $B(a)$ there is a term $t(a)$ such that we have

$$S_2^1 \vdash B(a) \to 1-IFRPr^*(t(a), \ulcorner B(\underset{\sim}{a})\urcorner). \quad \square$$

2. Induction-free proofs

The next proposition shows that induction can be efficiently eliminated from proofs of instances of T_2^i-provable formulas. The reader familiar with [6] can recognize it as an "arithmetic" version of the simulation of T_2^i by the propositions calculus G_i.

PROPOSITION 2.1

Let $A(a)$ be a #-free, sΣ_i^b-formula and assume:

$$T_2^i \vdash A(a).$$

Then for some term t we have:

$$S_2^1 \vdash \forall x,\ i\text{-IFRPr}(t(x), \ulcorner A(\underset{\sim}{x})\urcorner).$$

Proof

Let $A(a)$ satisfy the hypothesis of the proposition. By lemma 1.3 we have an LKB proof d of sequent:

$$2 \leqslant c_0, |a| \leqslant |c_0|, |c_0|\,|c_0| \leqslant |c_1|, \ldots, |c_0|\,|c_{k-1}| \leqslant |c_k| \to A(a),$$

such that d is #-free and consists only of sΣ_i^b-formulas.

By (meta)induction on the number of steps in d we show that whenever a sequent:

$$\Gamma(a, \bar{b}, \bar{c}) \to \Delta(a, \bar{b}, \bar{c})$$

occurs in d (with all free variables shown) then for any $n < \omega$ there is an i-IFR proof of the sequent:

$$a \leqslant \underline{n},\ b_0 \leqslant \underline{n}, \ldots, c_0 \leqslant \underline{n},\ \Gamma(a, \bar{b}, \bar{c}) \to \Delta(a, \bar{b}, \bar{c}). \qquad (*)$$

As everything will be effective (in n, proof d is fixed) the argument can be carried in S_2^1.

To avoid excessive notation we make the following simplifications. We shall not show explicitly the side formulas of the inference. All free variables other than eigenvariables of the inference of a sequent will be denoted $\bar{u}$; so these consist of a, $\bar{c}$ and all $\bar{b}$ with the exception of b_u, where b_u is the eigenvariable of the rule, and b_u itself will be denoted v (to avoid indices). The occurrences of $\bar{u}$ in formulas will not be explicitly indicated. Thus, for example, the sequent $(*)$ might be written as:

$$\bar{u},\, v \leqslant \underline{n},\, \Gamma(v) \to \Delta(v).$$

The only non-trivial cases in a step of the (meta)induction are the quantifier rules and the IND-rule. Let us consider these cases separately.

Case 1. Assume that the last inference was $\exists \leqslant$:right:

$$\frac{\to B(t)}{t \leqslant r \to (\exists x \leqslant r B(x))}$$

By (meta)induction assumption we have i-IFR proof of:

$$\bar{u},\, v \leqslant \underline{n} \to B(t).$$

An application of $\exists \leqslant$:right and a few exchanges gives:

$$\bar{u},\, v \leqslant \underline{n},\, t \leqslant r \to (\exists x \leqslant r B(x)).$$

Case 2. Assume that the last inference was $\forall \leqslant$:right:

$$\frac{v \leqslant r \to B(v)}{\to (\forall x \leqslant r B(x))}.$$

By (meta)induction assumption there is i-IFR proof of:

(1) $\bar{u},\, v \leqslant \underline{m},\, v \leqslant r \to B(v)$,

where m is minimal such that the sequent:

(2) $\bar{u} \leqslant \underline{n} \to r \leqslant \underline{m}$

has i-IFR proof (lemma 1.7). Such m exists of size $\leqslant n^{const}$, *const* depending on term r only.

Using (2), derive from (1) sequent:

$$\bar{u} \leqslant \underline{n},\, v \leqslant r \to B(v)$$

and by $\forall \leqslant$:right the required sequent:

$$\bar{u} \leqslant \underline{n} \to (\forall x \leqslant r B(x)).$$

Case 3. Rules ($\exists \leqslant$:left) and ($\forall \leqslant$:left) are dual to the right rules and are treated analogically.

Case 4. Assume that the last inference was IND:

$$\frac{B(v) \to B(s(v))}{B(0) \to B(r)}.$$

We may assume that term r does not contain variable v.

By the (meta)induction assumption there is an i-IFR proof d'_m of:

$$\bar{u},\ v \leqslant \underline{m},\ B(v) \rightarrow B(s(v)), \qquad (*)$$

where m is again chosen to be minimal such that sequent:

$$\bar{u},\ v \leqslant \underline{n} \rightarrow r \leqslant \underline{m}$$

has an i-IFR proof. Again $m \leqslant n^{const}$, i.e. $|m| = O(|n|)$.

We show how to construct from d'_m an i-IFR proof d_n of the required sequent:

$$\bar{u} \leqslant \underline{n},\ B(0) \rightarrow B(r).$$

This will be done in several steps: for $j = 0, 1, \ldots, |m|$ we successively construct i-IFR proofs D_j of sequents:

$$\bar{u},\ v \leqslant \underline{m},\ e \leqslant \underline{2}^j,\ v + e \leqslant \underline{m},\ B(v) \rightarrow B(v + e). \qquad (*)_j$$

Case $j = 0$ follows from d'_m, i.e. $(*)$, as there is an i-IFR proof d' of:

$$e \leqslant \underline{2}^0,\ B(v),\ B(s(v)) \rightarrow B(e + v).$$

As d' is constant for all n, it holds:

$$|D_0| = |d'_m| + O(|m|),$$

where $O(|m|)$ stands for a bound to the length of $(*)_0$ – this will be similar below as we are taking only $j \leqslant |m|$.

Assume we have constructed proof D_j of $(*)_j$. Then we construct D_{j+1} as follows.

(1) Using the substitution rule substitute $(v+f)$ for v in $(*)_j$ (f is new variable). This gives proof D_j^1 of:

$$\bar{u} \leqslant \underline{m},\ (v+f) \leqslant \underline{m},\ e \leqslant \underline{2}^j,\ (v+f) + e \leqslant \underline{m},\ B(v+f) \rightarrow B((v+f)+e) \qquad (*)'_j$$

of the length:

$$|D_j^1| = |D_j| + O(|m|).$$

(2) Substitute f for e in $(*)_j$ and apply cut with $(*)'_j$ to get proof D_j^2 of:

$$\bar{u} \leqslant \underline{m},\ v \leqslant \underline{m},\ f \leqslant \underline{2}^j,\ B(v),\ (v+f) \leqslant \underline{m},\ e \leqslant \underline{2}^j,$$
$$(v+f) + e \leqslant \underline{m} \rightarrow B((v+f)+e).$$

Obviously:

$$|D_j^2| = |D_j^1| + O(|m|)$$

(remember that D_j^1 is a prolongation of D_j so both sequents $(*)_j$ and $(*)'_j$ are already in D_j^1).

(3) Adding few weakenings and other trivial steps (using BASIC) produces proof D_n^3 of:

$$\bar{u} \leqslant \underline{m},\ v \leqslant \underline{m},\ v+(f+e) \leqslant \underline{m},\ e \leqslant \underline{2^j},\ f \leqslant \underline{2^j},\ B(v) \to B((v+f)+e).$$

As before:

$$|D_j^3| = |D_j^2| + O(|m|).$$

(4) There is a short i-IFR proof (constant for all n) of sequent:

$$f+e=g,\ B((v+f)+e) \to B(v+g).$$

This proof together with proof D_j^3 gives proof D_j^4 of:

$$\bar{u} \leqslant \underline{m},\ v \leqslant \underline{m},\ v+(f+e) \leqslant \underline{m},\ e \leqslant \underline{2^j},\ f \leqslant \underline{2^j},\ f+e=g,\ B(v) \to B(v+g$$

Again we have:

$$|D_j^4| = |D_j^3| + O(|m|).$$

(5) Obviously there is a proof of length $O(|m|)$ of:

$$v+g \leqslant \underline{m},\ g \leqslant \underline{2^{j+1}} \to$$
$$\to \left(\exists x \leqslant g\, \exists y \leqslant g\left((y+x)=g\ \&\ v+(y+x) \leqslant \underline{m}\ \&\ x \leqslant \underline{2^j}\ \&\ y \leqslant 2^j\right)\right),$$

which gives with D_j^4 proof D_j^5 of sequent:

$$\bar{u} \leqslant \underline{m},\ v \leqslant \underline{m},\ v+g \leqslant \underline{m},\ g \leqslant \underline{2^{j+1}},\ B(v) \to B(v+g).$$

As before:

$$|D_j^5| = |D_j^4| + O(|m|).$$

(6) Substituting e for g in the last sequent of (5) produces the required proof D_{j+1} of $(*)_{j+1}$. Obviously:

$$|D_{j+1}| = |D_j| + O(|m|) = |d'_m| + j \cdot O(|m|).$$

This completes the construction of proofs D_j.

Now we describe the construction of i-IFR proof d_n of the sequent:

$$\bar{u} \leqslant \underline{n},\ B(0) \to B(r).$$

(i) For $j = |m|$, D_j is an i-IFR proof of sequent:

$$\bar{u},\ v \leqslant \underline{m},\ e \leqslant \underline{m},\ v+e \leqslant \underline{m},\ B(v) \to B(v+e)$$

of length $|d'_m| + O(|m|^2)$, see (6) above.

(ii) By the choice of m, the following sequents obviously have i-IFR proofs:

$$\bar{u} \leqslant \underline{n} \to \bar{u} \leqslant \underline{m}, \qquad \bar{u} \leqslant \underline{n} \to r \leqslant \underline{m}.$$

(iii) From (i) and the first sequent of (ii) derive, substituting 0 for v:

$$\bar{u} \leqslant \underline{n},\ e \leqslant \underline{m},\ B(0) \to B(e).$$

(iv) Substituting now r for e in (iii) gives, together with (ii) and a few trivial inferences, the required sequent:

$$\bar{u} \leqslant \underline{n},\ B(0) \rightarrow B(r).$$

By (meta)induction assumption i-IFR proofs can be found of length:

$$|d'_m| \quad |m|^{O(1)}.$$

As $|m| = |n|^{O(1)}$ we get (cf. (i)):

$$|d_n| = |n|^{O(1)}.$$

This proves case 4.

Now we are ready to prove the proposition; we use the original names for variables from (∗) now.

Take the constructed i-IFR proof d_n of:

$$a,\ \bar{c} \leqslant \underline{n},\ 2 \leqslant c_0,\ |a| \leqslant |c_0|, \ldots, |c_0|\,|c_{k-1}| \leqslant |c_k| \rightarrow A(a).$$

Substitute for a numeral $\underline{n}$ and for c_l, $l = 0, \ldots, k$, numerals $\underline{n_l}$ making the antecedent true and by lemma 1.7 i-IFR-provable. As k is a fixed, standard number this can be performed in S_2^1.

It remains to show that the constructed proof can be augmented by additional terms $\bar{t}$ and supplementary proofs $\bar{d}'$, as is required in definition 1.4. First observe that the quantifier nesting (i.e. the number of quantifier inferences and their order) is essentially the same as in the original (standard) proof of $A(a)$. Then define:

$$t_i(\bar{a}) = r_i(\bar{a},\ b_0/t_0(\bar{a}), \ldots, b_{i-1}/t_{i-1}(\bar{a})),$$

where $r_i(\bar{a}, b_0, \ldots, b_{i-1})$ is the term in the elimination rule as in definition 1.4, clause (iv). The supplementary proofs are easily constructed.

Finally, the whole construction is effective (in n) with polynomial bounds so the whole argument can be carried in S_2^1. □

The significance of the proposition is that the constructed proofs have bounded complexity. Without this requirement one can get short proofs of instances $A(\underline{n})$ e.g. by the methods of definable cuts, cf. [9].

The difference between T_2^i and S_2^i is that in the case of T_2^i proofs of instances $A(\underline{n})$ are generally only sequence-like while in the case of S_2^i tree-like proofs can be constructed. We state the following statement which is well-known, cf. [1, chapter 4].

PROPOSITION 2.2

Let $A(a)$ be a #-free, $s\Sigma_i^b$-formula and assume:

$$S_2^i \vdash A(a).$$

Then there is a term t such that:

$$S_2^i \vdash \forall x,\ i\text{-IFRPr}^*(t(x), \ulcorner A(\underset{\sim}{x}) \urcorner).$$

Proof

The proof goes as in the preceding proposition, the only difference being the treatment of PIND which is rather straightforward: from

$$B\left(\left\lfloor \frac{b}{2} \right\rfloor\right) \rightarrow B(b)$$

we get (by repeating the proof $|n|$ times) $|n|$ proofs of:

$$B(\underline{n}_j) \rightarrow B(\underline{n}_{j+1}),$$

where $n_{|n|} := n$ and $n_j := \lfloor n_{j+1}/2 \rfloor$, i.e. $n_0 = 0$. Joining these proofs by cuts we get the required proof of:

$$B(0) \rightarrow B(\underline{n}). \quad \square$$

3. Truth definition and witnessing function

We shall need certain partial truth definitions for #-free, $s\Sigma_i^b$-formulas. This material is rather familiar, see [8] or any of [3–5,10], and thus our exposition is brief.

There is a term *val(x, y)* such that for any #-free term $t(\bar{a})$ and any evaluation $\overline{m}$ of free variables $\bar{a}$ we have:

$$t(\bar{a}/\overline{m}) \leqslant val(\ulcorner t(\bar{a}) \urcorner, \langle \overline{m} \rangle).$$

Here $\langle \overline{m} \rangle$ is a code of sequence $\overline{m}$ and $\ulcorner t(a) \urcorner$ the Gödel number of $t(\bar{a})$. (We can take for *val(x, y)* roughly $y\#x$.) It follows that there is a Δ_1^b w.r.t. S_2^1 definition of the value of a #-free term and thus there is also a Δ_1^b w.r.t. S_2^1 partial truth definition for quantifier free formulas. Using such a partial truth definition it is routine to write down a partial truth definition TR_i, a Σ_i^b-formula, satisfying Tarski's conditions for #-free, $s\Sigma_i^b$-formulas. In particular,

$$S_2^i \vdash \forall \bar{x},\ A(\bar{x}) \equiv \mathrm{TR}_i(\ulcorner A \urcorner, \langle \bar{x} \rangle)$$

holds for every #-free, $s\Sigma_i^b$-formula A.

DEFINITION 3.1

For $j \leqslant i$, j-RFN(i-R) is the following formula:

$$\text{“}\forall A \in s\Sigma_j^b \forall x, \bar{y}\big(i\text{-RPr}(x, A(a)) \rightarrow \mathrm{TR}_j(A, \bar{y}\big).$$

Formulas j-RFN(i-IFR), j-RFN(i-R*) and j-RFN(i-IFR*) are defined analogically using i-IFRPr resp. i-RPr* resp. i-IFRPr* instead of i-RPr in the above definition. $\square$

The formulas defined above formalize the reflection principles for the i-regular provability notion with or without the requirement to be induction-free or in a tree-form.

PROPOSITION 3.2

For $i \geqslant 1$ it holds:

$$T_2^i \vdash i\text{-RFN}(i\text{-R}).$$

Thus T_2^i also proves formulas: i-RFN(i-R*), i-RFN(i-IFR) and i-RFN(i-IFR*).

Proof

Work in T_2^i and assume $D = \langle d, \bar{t}, \bar{d}' \rangle$ is an i-regular proof of formula $A(\bar{a})$ which is $s\Sigma_i^b$. Let d be the sequence of sequents $S_1(\bar{a}, \bar{b}), \ldots, S_r(\bar{a}, \bar{b})$ where $\bar{b}$ are all other free variables of d, like in definition 1.4. Thus S_r is $\rightarrow A(\bar{a})$. By induction on $p \leqslant r$ show:

$$\neg \mathrm{TR}_i(A, \bar{u}) \rightarrow \exists \bar{v} \leqslant w(\bar{u}) \exists q \leqslant r,$$

$$\text{“}S_q(\bar{u}, \bar{v}) \text{ is not true”}\big) \ \& \ \big(q \leqslant r - p \text{ or “}S_q \text{ is initial”}\big)\text{”}.$$

This formula is easily written using TR_i and it is Σ_{i+1}^b. Hence it is provable in S_2^{i+1} and thus also in T_2^i, cf. the conservation result of [2]. The bounds w to $\bar{v}$ are obtained from the additional terms $\bar{t}$ guaranteed by D and their correctness is verified using the supplementary proofs $\bar{d}'$ (for details of such an argument see [4,5,10]).

Taking $p := r - 1$ and observing that each initial sequent must be true establishes the statement. □

In the following T_k^i is a theory defined as T_2^i but having function symbol $\#_k$ (and appropriate axioms in BASIC) instead of $\#$.

Function $\#_k$ is:

$$x \#_1 y := x \cdot y,$$

$$x \#_{k+1} y := 2^{|x| \#_k |y|}.$$

Hence $\#_2$ is $\#$.

COROLLARY 3.3 [1]

For $i, k \geqslant 1$ we have:

(a) T_{k+1}^i is not $\forall \Pi_1^b$-conservative over T_k^i.

(b) S_{k+2}^{i+1} is not $\forall \Pi_1^b$-conservative over S_{k+1}^{i+1}.

[1] This corollary was also obtained by P. Pudlák.

Proof

(a) From proposition 3.2 it follows, in particular, that T_2^i proves a consistency statement about T_1^i:

$$\forall x, \neg i\text{-RPr}(x, \ulcorner \rightarrow \urcorner).$$

On the other hand T_1^i does not prove it, see [4].

For general k we apply an idea from [5]. Fix $k > 1$. Say that d is a *restricted* T_k^i proof of a #-free $s\Sigma_i^b$-formula $A(a)$ iff d is an i-regular proof of a sequent of the form:

$$2 \leqslant |c|^{(k)}, |a|^{(k)} \cdot |a|^{(k)} \cdot \ldots \cdot |a|^{(k)} \leqslant |c|^{(k)} \rightarrow A(a), \qquad (*)$$

where $|a|^{(k)}$ appears j times *and* $j \leqslant |d|^{(k+1)}$. (Here $|a|^{(k)}$ is a k times iterated function $|x|$ on a.) Although j grows slowly with d it can exceed every standard number. Therefore similarly as in lemma 1.3 (by compactness): a #-free $s\Sigma_i^b$-formula is provable in T_k^i iff it has a restricted T_k^i proof.

Given a and d, a number c of size about $a\#_{k+1}d$ can be found to make the antecedent of $(*)$ true. As T_2^i proves reflexiveness for i-regular proofs we conclude that T_{k+1}^i proves that every formula A with restricted T_k^i proof is correct.

Now take for $A(a)$ a diagonal, #-free $s\Pi_1^b$-formula s.t.:

$$S_2^1 \vdash \forall x[A(x) \equiv \forall d \leqslant x(\text{“}d \text{ is not a restricted } T_k^i \text{ proof of } A(a)\text{”})].$$

Clearly $T_k^i \nvdash A(a)$, but by the argument above $T_{k+1}^i \vdash A(a)$.

(b) This follows from (a) and because for $l \geqslant 2$:

$$S_l^{i+1} \succsim_{\forall \Pi_1^b} T_l^i,$$

which is an easy consequence of the conservation result from [2]. □

Next we relate i-R provability to propositional provability in G_i:

COROLLARY 3.4

For $1 \leqslant j \leqslant i$ the following formulas are equivalent in S_2^1: i-RFN(G_j), RFN(i-R) and j-RFN(i-IFR).

Also formulas Con(G_i), Con(i-R) $:= \forall x, \neg i\text{-RPr}(x, \ulcorner \rightarrow \urcorner)$ and Con(i-IFR) $:= \forall x, \neg i\text{-IFR}(x, \ulcorner \rightarrow \urcorner)$ are equivalent in S_2^1.

Remark

It follows that there are polynomial time functions f, g (definable in S_2^1) which assign to an i-R proof d an i-IFR proof $f(d)$ of the same end-sequent and a G_i proof $g(d)$ of its propositional translation (and vice versa).

For the definition of G_i, j-RFN(G_i) and Con(G_i) the reader should consult [6].

Proof of corollary 3.4

We prove the case of the consistency statements; the argument for the reflection principles is identical.

By [6] Con(G_i) is the strongest $\forall\Pi_1^b$-formula provable in T_2^i (over S_2^1). By proposition 3.2 (as i-RFN(i-R) implies trivially Con(i-R)) formula Con(i-R) is provable in T_2^i. Obviously (in S_2^1) Con(i-R) implies Con(i-IFR), we have:

$$S_2^1 \vdash \mathrm{Con}(G_i) \to \mathrm{Con}(i\text{-R}),$$

$$S_2^1 \vdash \mathrm{Con}(i\text{-R}) \to \mathrm{Con}(i\text{-IFR}).$$

Hence to close the argument it is sufficient to derive (in S_2^1) Con(G_i) from Con(i-IFR).

By proposition 2.1 we have:

$$S_2^1 \vdash \forall x,\ i\text{-IFRPr}\big(t_0(x),\ \ulcorner \neg \mathrm{Pr}f_{G_i}(\underset{\sim}{x},\ \ulcorner \to \urcorner)\urcorner\big),$$

where $\mathrm{Pr}f_{G_i}$ is a Δ_1^b-formalization of provability in G_i; $\mathrm{Pr}f_{G_i}$ can be taken as a subformula of UNIV. By lemma 1.7 we have then too:

$$S_2^1 \vdash \forall x,\ \mathrm{Pr}f_{G_i}(x,\ \ulcorner \to \urcorner) \to i\text{-IFRPr}\big(t_1(x),\ \ulcorner \mathrm{Pr}f_{G_i}(\underset{\sim}{x},\ \ulcorner \to \urcorner)\urcorner\big).$$

Putting this together gives (as i-IFR proofs are provably closed under cuts):

$$S_2^1 \vdash \forall x,\ \mathrm{Pr}f_{G_i}(x,\ \ulcorner \to \urcorner) \to i\text{-IFRPr}(t_2(x),\ \ulcorner \to \urcorner).$$

This is required:

$$S_2^1 \vdash \mathrm{Con}(i\text{-IFr}) \to \mathrm{Con}(G_i). \quad \square$$

For obtaining the conservation result $T_2^i \succcurlyeq_{\Sigma_i^b} S_2^i$ it would be enough to prove the reflection principle i-RFN(i-IFR) in S_2^i. We are not able to do this, neither are we able to show the independence of this formula from S_2^i. However, it turned out that the main obstacle is not the quantifier complexity but the structure of the proof-figures: whether the proofs are or are not in a tree form. We have the following statement.

PROPOSITION 3.5

For $i \geqslant 1$ it holds:

$$S_2^i \vdash i\text{-RFN}(i\text{-IFR}^*).$$

Proof

The idea of the proof is the same as before; by induction on the number of sequents in an i-IFR* proof show that all sequents in it are true. However, as we now work in S_2^i instead of S_2^{i+1} we must decrease the complexity of the assertion that a sequent is true. This is provided by formalizing the witnessing theorem of [1] in S_2^i.

Using formula TR_{i-1} we can construct a partial truth definition $\overline{\mathrm{TR}}_{i-1}$ for $s\Sigma^b_{i-1} \cup s\Pi^b_{i-1}$-formulas. $\overline{\mathrm{TR}}_{i-1}$ is Δ^b_i w.r.t. S^1_2. With $\overline{\mathrm{TR}}_{i-1}$ in hands we can formalize the witness formula of [1] for $s\Sigma^b_i$-formulas in the following way:

$\mathrm{Witness}^{i,\bar{a}}_A(\bar{a}, w)$

is the formula:

"if $A \in s\Sigma^b_{i-1} \cup s\Pi^b_{i-1}$ then

$\overline{TR}_{i-1}(A, \langle \bar{a} \rangle)$, and

if A is of the form $\exists x \leqslant t(\bar{a}) B(\bar{a}, x)$ and $A \in s\Sigma^b_i \setminus s\Sigma^b_{i-1}$ then

$\overline{TR}_{i-1}(b \leqslant t(\bar{a}) \wedge B(\bar{a}, b), \langle \bar{a}, w \rangle)$."

For $\Gamma = (A_1, \ldots, A_j)$ a cedent of $s\Sigma^b_i$-formulas, analogically with [1] define:

$\mathrm{Witness}^{i,\bar{a}}_{\wedge\Gamma}(\bar{a}, w)$

is the formula:

$\forall j \leqslant |\Gamma|, \mathrm{Witness}^{i,\bar{a}}_{A_j}(\beta(j, w), \bar{a})$,

and

$\mathrm{Witness}^{i,\bar{a}}_{\vee\Gamma}(\bar{a}, w)$

is the formula

$\exists j \leqslant |\Gamma|, \mathrm{Witness}^{i,\bar{a}}_{A_j}(\beta(j, u), \bar{a})$.

Here β is the standard coding function, cf. [1]. Observe that all these witness formulas are also Δ^b_i w.r.t. S^1_2.

Let us look now how witnessing functions for sequents in an i-IFR* proof $D = \langle d, \bar{t}, \overline{d'} \rangle$ are constructed. By the definition of $s\Sigma^b_i$-formulas, the only rules which can have as the principal formula an $s\Sigma^b_i$-formula not in $s\Sigma^b_{i-1}$ are: $\exists \leqslant$ rules, contractions, exchanges and cut-rule.

In $\exists \leqslant$:left a new witness function is obtained essentially by renaming a variable in the witness function for the upper sequent.

In $\exists \leqslant$:right one computes a witness of the principal formula by evaluating a #-free term.

In contraction :left two variables are given the same name.

In contraction :right no new values are computed but a $s\Pi^b_{i-1}$ oracle (the kernel of the principal formula) is queried which witness is the correct one.

In exchange rules only variables or values are permuted.

Finally, in cut-rule first a witness for the succedent containing the cut formula is computed and then substituted for variables in the computation of a witness for the succedent of the other sequent.

Thus if $\eta(S)$ is the number of #-free terms evaluated during the computation (as above) of the witness for a sequent S we have:

$$\eta(S) \leqslant 1 + \eta(S_0),$$

in case of all rules above except cut-rule, S_0 being the upper sequent, and:

$$\eta(S) \leqslant \eta(S_0) + \eta(S_1)$$

in case of cut-rule, S_0 resp. S_1 being the upper sequents.

Let $\xi(S)$ be the number of sequents in d above S. Then using in an essential way the assumption that d is in a tree-form we have:

$$\eta(S) \leqslant \xi(S) \leqslant |d| \leqslant |D|. \qquad (+)$$

Moreover, bounds in outermost $\exists$-quantifiers in $s\Sigma_i^b$-formulas in the succedent of a sequent give a priori bounds to values of terms used in the computation of the witness. Let $r(\bar{a}, \bar{b})$ be the greatest of these bounds.

We shall be interested only in computing witnesses with non-parametrical free variables $\bar{b}$ in d bounded by $\bar{t}$ of D, i.e.: $b_i \leqslant t_i(\bar{a})$, as in definition 1.4. Then values of $r(\bar{a}, \bar{b})$ are bounded by:

$$r(\bar{a}, b_j/t_j(\bar{a})). \qquad (++)$$

Using the fact that terms r and $\bar{t}$ are part of D and that for a #-free term $s(\bar{a})$ it holds in general:

$$val(s(\bar{a})) \leqslant \max(\bar{a}, 2)^{|s|},$$

$|s|$ being the length of term s, we can replace bound $(++)$, by:

$$\max(\bar{a}, 2)^{|D|^2}. \qquad (+++)$$

Bounds $(+)$ and $(+++)$ imply that under the assumption $b_i \leqslant t_i(\bar{a})$ the whole computation of a witness described above requires evaluation of at most $|D|$ #-free terms with values at most $\max(\bar{a}, 2)^{|D|^2}$, and so the computation itself can be coded below:

$$\max(\bar{a}, 2)^{O(|D|^3)}. \qquad (*)$$

It is now routine to write down a Σ_i^b-definition (Δ_i^b w.r.t. S_2^i, in fact) of the function

$$F(D, S, \bar{a}, \bar{b}, w) = v$$

satisfying:

$$\forall b_j \leqslant t_j(\bar{a}), \text{Witness}^{i,\bar{a},\bar{b}}_{\wedge\Gamma}(\bar{a}, \bar{b}, w) \to \text{Witness}^{i,\bar{a},\bar{b}}_{\vee\Delta}(\bar{a}, \bar{b}, v), \qquad (**)$$

for S a sequent $\Gamma \to \Delta$ in any i-IFR* proof D.

Function F is defined by induction on the number of sequents in D above S, considering several clauses as in the above discussion. The explicit bounds $(*)$ to the (code of) computations guarantees (by Σ_i^b-PIND) that the computations are defined and output some values.

Formula $(**)$ is Π_i^b, it is obviously true for S being an initial sequent and its validity for the upper sequents of an inference implies its validity for the lower sequent too. Thus by Π_i^b-PIND $(**)$ holds for the end-sequent of D as well. The end-sequent S_e has the form:

$$\rightarrow A(\bar{a}),$$

$A \in s\Sigma_i^b$, and so it holds:

$$\text{Witness}_A^{i,\bar{a}}\big(\bar{a}, F\big(D, S_e, \bar{a}, \bar{0}, 0\big)\big). \qquad (***)$$

Now it is provable in S_2^i, cf. [1, chapter 5], that:

$$\text{Witness}_A^{i,\bar{a}}(\bar{a}, v) \rightarrow \text{TR}_i(A, \langle\bar{a}\rangle).$$

Hence $(***)$ yields:

$$i\text{-IFR}^*(u, A) \rightarrow \text{TR}_i(A, \bar{a}),$$

that is:

$$i\text{-RFN}(i\text{-IFR}^*).$$

This completes the proof of proposition 3.5. □

The crucial use of the assumption that the proof is in a tree-form is in the derivation of bound $(+)$. In general one gets only:

$$\eta(S) \leqslant 2^{\xi(S)} \leqslant D,$$

which is not good enough as in the later bounds $\eta(S)$ occurs in an exponent.

4. Proofs in a tree-form

Proposition 3.5 shows that the Σ_i^b-conservativeness of T_2^i over S_2^i would follow if we could (in S_2^1) put i-IFR proofs (or G_i proofs) into a tree-form and enlarge their length only polynomially. The opposite implication is also true.

PROPOSITION 4.1

For $i \geqslant 1$, T_2^i is $\forall\Sigma_i^b$-conservative over S_2^i iff the following formulas are provable in S_2^1:

(a) $\forall x$, sentence $y \exists z$, $i\text{-IFRPr}(x, y) \rightarrow i\text{-IFRPr}^*(z, y)$,

(b) TREE(G_i).

((a) and (b) are equivalent over S_2^1.)

Proof

(a) The "if part" follows from proposition 3.5 as by corollary 3.4 and [6] formula i-RFN(i-IFR) is the strongest (over S_2^1) $\forall\Sigma_i^b$-formula provable in T_2^i, and is obviously implied by (a).

Now assume $T_2^i \succcurlyeq_{\Sigma_i^b} S_2^i$ and thus:

$S_2^i \vdash i\text{-RFN}(i\text{-IFR})$.

In particular (y stands here and below for an $s\Sigma_i^b$-sentence):

(1) $S_2^i \vdash i\text{-IFRPr}(x, y) \to \text{TR}_i(y, 0)$.

By lemma 1.7 we have:

(2) $S_2^1 \vdash i\text{-IFRPr}(x, y) \to 1\text{-IFRPr}^*\big(t_0(x, y), \ulcorner i\text{-IFRPr}(\underset{\sim}{x}, y)\urcorner\big)$,

for some term t_0.

Applying proposition 2.2 to (1) gives for some term t_1:

(3) $S_2^1 \vdash i\text{-IFRPr}^*\big(t_1(x, y), \ulcorner i\text{-IFRPr}(\underset{\sim}{x}, \underset{\sim}{y}) \to \text{TR}_i(\underset{\sim}{y}, 0)\urcorner\big)$.

Clauses (2) and (3) readily give:

(4) $S_2^1 \vdash i\text{-IFRPr}(x, y) \to i\text{-IFRPr}^*\big(t_2(x, y), \ulcorner \text{TR}_i(\underset{\sim}{y}, 0)\urcorner\big)$,

for some term t_2.

Tarski's conditions for TR_i are proved by the complexity of sentence y and the proof is in a tree-form. Hence we have:

(5) $S_2^1 \vdash i\text{-IFRPr}^*\big(t_3(x, y), \ulcorner \text{TR}_i(\underset{\sim}{y}, 0)\urcorner \to y\big)$,

t_3 a term.

Finally, (4) with (5) gives (for t_4 a term):

(6) $S_2^1 \vdash i\text{-IFR}(x, y) \to i\text{-IFR}^*(t_4(x, y), y)$.

This proves the proposition.

The statement (b) for G_i follows as G_i-provability is equivalent to i-IFR provability. Note that there is also a direct proof for G_i following the lines above and using propositional versions of (2), (3) and (5). □

Similarly with [6] where the Σ_j^b-conservativeness of T_2^i over S_2^i, $j < i$, was characterized as essentially a combinatorial question concerning polynomial simulations, proposition 4.1 offers a reformulation of Σ_i^b-conservativity in terms of the efficiency of the sequence-form versus the tree-form of G_i proofs (resp. i-IFR proofs).

Our results also imply that G_i^* – a propositional proof system defined as G_i but with proofs only in a tree-form – has the same relation to S_2^i as G_i to T_2^i, cf. [6].

Acknowledgement

We thank S. Buss for a discussion on the manuscript of this paper and two anonymous referees for their suggestions about its presentation.

References

[1] S. Buss, *Bounded Arithmetic* (Bibliopolis, Napoli, 1986).

[2] S. Buss, Axiomatization and conservation results for fragments of bounded arithmetic, to appear in: *Contemporary Mathematics AMS vol. 106, Proc. Workshop in Logic and Computation* (1990) pp. 57–84

[3] H. Gaifman and C. Dimitracopoulos, Fragments of Peano's arithmetic and the MRDP theorem, *Logic et Algorithmic*, Monogr. No. 30 de l'Enseignement Mathématique, Genève (1982) pp. 187–206.

[4] J. Krajíček, π_1-conservativeness in systems of bounded arithmetic, typescript (1988).

[5] J. Krajíček, Exponentiation and second order bounded arithmetic, Ann. Pure Appl. Logic 48 (1990) 261–276.

[6] J. Krajíček and P. Pudlák, Quantified propositional calculi and fragments of bounded arithmetic, Zeit. Math. Logik and Grandlog. 36(1) (1990) 29–46.

[7] J. Krajíček, P. Pudlák and G. Takeuti, Bounded arithmetic and the polynomial hierarchy, Ann. Pure Appl. Logic 52 (1991) 143–153.

[8] J. Paris and C. Dimitracopoulos, Truth definition for Δ_0 formulae, *Logic et Algorithmic*, Monogr. No. 30 de l'Enseignement Mathématique, Genéve (1982) pp. 317–330.

[9] J. Paris and A. Wilkie, On the scheme of induction for bounded arithmetic formulas, Ann. Pure Appl. Logic 35(3) (1987) 261–302.

[10] G. Takeuti, Bounded arithmetic and truth definition, Ann. Pure Appl. Logic 39(1) (1988) 75–104.