Implicit proofs

Jan Krajíček*

Mathematical Institute†
Academy of Sciences, Prague

Abstract

We describe a general method how to construct from a propositional proof system P a possibly much stronger proof system iP. The system iP operates with exponentially long P-proofs described "implicitly" by polynomial size circuits.

As an example we prove that proof system iEF, *implicit* EF, corresponds to bounded arithmetic theory V_2^1 and hence, in particular, polynomially simulates the quantified propositional calculus G and the Π_1^b-consequences of S_2^1 proved with one use of exponentiation. Furthermore, the soundness of iEF is not provable in S_2^1. An iteration of the construction yields a proof system corresponding to $T_2 + Exp$ and, in principle, to much stronger theories.

Extended Frege system EF is considered to be a strong propositional proof system. The qualification strong means that EF smoothly formalizes many arguments in elementary combinatorics or algebra and it seems very hard to come up with tautologies that would be hard to prove in EF (i.e. that they would require long proofs). Another strong proof system is the quantified propositional calculus G which operates with quantified propositional formulas. We can move up in this hierarchy allowing a proof system to quantify also over boolean functions, functionals, etc. But besides simulating definitions from higher order arithmetic or set theory we do not really have any other way of directly constructing strong proof systems.

The qualification directly is important here as we do have a general correspondence between proof systems and first-order theories (obeying certain

*Partially supported by grant # A 101 99 01 of the Academy of Sciences of the Czech Republic and by project LN00A056 of The Ministry of Education of the Czech Republic.

†Also member of the Institute for Theoretical Computer Science of the Charles University. A part of this work was done while visiting the Mathematical Institute at Oxford.

tame technical conditions satisfied by all "usual" theories, including set theory) and, in particular, we can define a strong proof system from a strong theory. This correspondence is very useful and it is the deepest information applying to all proof systems (as opposed to statements about particular ones) that we have. In particular, the statements above that EF and G are strong could be substantiated by identifying theories corresponding to them (S^1_2 and U^1_2, respectively; see the references given below). (The proof system extending G by allowing the quantification over functions, functionals, etc. corresponds to $T_2 + Exp$ or to a bit stronger theory, depending on the exact definition).

However, our aim here is to investigate a possibility of a direct, essentially combinatorial, description of strong proof systems that would, in particular, not refer to first order theories. This appears of interest in connections with several problems (e.g. a combinatorial characterization of hard tautologies and of consistency statements in particular, the existence of an optimal proof system, constructions of models of strong bounded arithmetic theories, etc.).

In proof complexity there are several interesting results of the form of an upper bound on the size of proofs of particular formulas or of the form of a polynomial simulation of one proof system by another. These results tend to be much simpler to prove using bounded arithmetic than using direct proof manipulations. Thus although we want to bypass the reference to theories in definitions of strong proof systems, we shall use the correspondence between proof systems and theories in proofs. However, the concept of implicit EF (and iP in general) is defined without any reference to arithmetic.

Let us now describe a part of this correspondence that we will need (and fix the notation in the process). A $\forall\Pi^b_1$-sentence $\forall x, \psi(x)$, with $\psi(x)$ having the form $\forall y(|y| \leq |x|^{O(1)}), \psi_0(x, y)$ for some p-time predicate ψ_0, determines an infinite sequence of propositional formulas $||\psi(x)||^n$ as follows. The formula has n atoms $p_1, \dots, p_n$ for bits of an x, some $n^{O(1)}$ atoms $q_1, \dots, q_m$ for bits of a y in ψ_0, and further it has $n^{O(1)}$ atoms $r_1, \dots, r_s$ for bits of values of subcircuits of a fixed (canonically constructed) circuit computing from $\overline{p}$, $\overline{q}$ the truth value of $\psi_0(x, y)$. The formula $||\psi(x)||^n$ expresses in a DNF form that if $\overline{r}$ are correctly computed by the circuit from the inputs $\overline{p}$, $\overline{q}$ then the output of the computation is 1. A number b of length n is identified with a binary string $(b_1, \dots, b_n)$ of length n, and these bits will make $||\psi(x)||^n(p_i/b_i)$ a tautology iff $\psi(b)$ is true.

The correspondence between a theory T and a proof system P implies, in particular, the following:

- If T proves $\forall x; \psi(x)$ then tautologies $||\psi(x)||^n(b)$ have polynomial size

P-proofs.

- T proves the soundness of P and for any another proof system Q, if T proves also the soundness of Q then P polynomially simulates Q.

This correspondence has been discovered by Cook [2] (he considered the key case of $T = PV$ and $P = EF$). The two properties of the correspondence between S^1_2 and EF has been proved by Buss [1], between U^1_2 and G by Krajíček and Takeuti [9], and the case of general T and P was treated in Krajíček and Pudlák [7].

We shall not repeat other definitions and basic facts from proof complexity or bounded arithmetic. The reader can find those in [5] (or in the other original references listed in the bibliography).

1 Implicit EF

A proof system is a polynomial-time function P whose range is exactly the set $TAUT$ of tautologies in the DeMorgan language, cf.[3]. A P-proof of τ is any string π such that $P(\pi) = \tau$. The idea of implicit proofs is that instead of representing π of length ℓ by writing down it's bits $\pi_1, \dots, \pi_\ell$ we present a circuit β with $\log(\ell)$ inputs that computes π_i from $i \leq \ell$. The advantage of this implicit description of π is that β can be, in principle, exponentially smaller than π. However, the circuit β alone does not constitute a proof of anything. In order to get a proof system we supplement β with an ordinary P-proof α of the fact that β indeed describes a valid P-proof.

We will consider this general definition in Section 3. Now we will confine ourselves to EF. This particular case allows to achieve a full generality of the construction while having a nice intuitive property of β: The circuit computes whole formulas forming the steps of an EF-proof rather than just individual bits. This is useful in developing the connections with bounded arithmetic. We show in Section 3 that even with this property of β nothing is lost in generality.

Let EF be a fixed Extended Frege system in the DeMorgan language. The set of all DeMorgan tautologies is denoted $TAUT$. We shall assume that EF proofs are written in an *enhanced* form where each step caries an information about the rule and the previous steps that were used in its derivation. This is an inessential change that does not affect the proof complexity of EF (more than by a logarithmic factor).

The symbol $\leq_{\text{lex}}$ denotes the lexicographic ordering on any fixed $\{0,1\}^k$. If we identify $i = (i_1, \dots, i_k) \in \{0,1\}^k$ with the number $\sum_{j: i_j \neq 0} 2^j$ then $\leq_{\text{lex}}$

corresponds to the usual ordering on $\{0, \ldots, 2^k - 1\}$.

Definition 1.1 *Let $\tau \in TAUT$. An* implicit *EF proof of τ is a pair (α, β) such that:*

1. *β is a many-output boolean circuit in variables $i_1, \ldots, i_k$.*
2. *The sequence $\beta(\overline{0}), \ldots, \beta(i), \ldots, \beta(\overline{1})$ is an EF-proof of τ (the i's are ordered by $\leq_{lex}$).*
   The EF-proof described by β is denoted β^*.
3. *α is an EF-proof of a (canonical) tautology $Correct_\beta(x_1, \ldots, x_k)$ expressing that*

   "the formula in the step $\beta(x_1, \ldots, x_k)$ has been derived

   in β^* according to the EF-rules specified in $\beta(x_1, \ldots, x_k)$"

The proof system so defined is denoted iEF.

Note that we do not need to require that α also contains an EF-proof of the fact that the last step of β^* is τ (plus the auxiliary information); that is expressed by a true boolean sentence written using a circuit and so it always has a polynomial size proof in EF. Further note that as we consider enhanced EF-proofs the formula $Correct_\beta(\overline{x})$ is indeed expressible without existential quantification over steps in β^*, and hence if β^* is a correct EF-proof the formula is a tautology (when considering only polynomial size proofs such a quantification posses no problem as the quantifiers range only over a polynomial size set). The size of $Correct_\beta$ is $O(|\beta|)$.

The formulas in β^* are of the size at most $|\beta|$ while their number can be up to $2^{\Omega(|\beta|)}$. This would pose an apriori restriction for a proof system like a Frege system. However, for EF this is not a restriction due to the presence of the Extension rule, as we shall see in the proof of Theorem 2.1.

Let us start with the obvious.

Lemma 1.2 *iEF is a proof system in the sense of Cook-Reckhow [3], and it polynomially simulates EF.*

Proof :

It is clear that iEF is sound and complete. The third condition in the Cook-Reckhow's definition is that the relation "(α, β) is an iEF -proof of τ" is decidable in polynomial time. That follows as it is sufficient to check that the formula in the last step of β^* is τ, and that "α is an EF-proof of $Correct_\beta$" which is a polynomial time relation obviously.

A p-simulation of EF by iEF proceeds as follows. Let π be an EF-proof of τ of size m. Let β be a circuit in $\log(m)$ inputs that simply copies π into β^*, i.e. $\beta^* = \pi$. Clearly such β exists of size $O(|\pi|)$.

For α we take an EF-proof of $||Prf(u,v)||^m(\pi,\tau)$, where $Prf(u,v)$ is the polynomial time relation "u is an EF-proof of v". This has an EF-proof of size $O(|\pi|^2)$ that is constructed by a polynomial time algorithm from π and τ. This completes the p-simulation.

q.e.d.

Another p-simulation of EF by iEF follows from Lemma 4.1.

2 The strength of iEF

Now we calibrate the strength of iEF.

Theorem 2.1 *iEF corresponds to bounded arithmetic theory V_2^1. In particular,*

1. *V_2^1 proves the soundness of iEF.*
2. *Whenever a $\forall\Pi_1^b$-sentence $\forall x\psi(x)$ is provable in V_2^1 then the sequence of tautologies $||\psi(x)||^n$ has polynomial size iEF-proofs.*
3. *If V_2^1 proves the soundness of a proof system Q then iEF polynomially simulates Q.*

Moreover, an iEF-proof of $||\psi(x)||^n$ can be constructed by a polynomial-time algorithm (from a string of length n) and the construction can be formalized in S_2^1, and the polynomial simulation in item 3. can be also defined in S_2^1.

Proof :

We start by proving the soundness of iEF in V_2^1. Work in a model of V_2^1 where we have an iEF-proof (α, β) (coded by a number, say b) of formula τ. Let a be a number coding a truth assignment to atoms of τ.

By induction on $\overline{i} \in \{0,1\}^k$ (ordered by $\leq_{\text{lex}}$) construct a set $A_{\overline{i}}$ coding a truth assignment to extension atoms in β^* introduced in steps $\leq_{\text{lex}}\overline{i}$ such that all their extension axioms are true when atoms of τ are evaluated by a. The induction step is trivial and the statement that such a set exists is

$\Sigma_1^{1,b}$, hence the $\Sigma_1^{1,b}$-induction implies that there is such a set $A := A_{\overline{1}}$ for $\overline{i} = \overline{1}$.

Using A, a and b as parameters prove by Π_1^b-induction on $\overline{i}$ that all formulas in β^* are true under the assignment given by a and A. The induction step uses the proof α: EF is sound in any model of V_2^1 and hence each step of β^* is indeed derived correctly via EF-rules, which are all sound. Hence τ is satisfied by (any) assignment a. This completes the proof of the first part.

Assume that V_2^1 proves a $\forall\Pi_1^b$-sentence $\forall x, \psi(x)$ which is of the form $\forall y \psi_0(x, y)$ with y implicitly bounded in a Δ_1^b-formula ψ_0. We shall describe polynomial size iEF-proofs of tautologies $||\psi(x)||^n$, $n \geq 1$. In fact, the proof π_n of $||\psi(x)||^n$ will be constructed by a polynomial time algorithm from a string of length n, and the construction itself could be formalized in S_2^1.

By [4] the hypothesis implies (is equivalent to, in fact) that there is a term $t(x)$ of the language of S_2^1 such that S_2^1 proves:

$$(*) \qquad\qquad t(x, y) \leq |z| \longrightarrow \psi_0(x, y) \ .$$

Furthermore, we may assume that $(*)$ has an S_2^1-proof in which all formulas are strict Σ_1^b; let Ω be one such proof. The algorithm that will construct π_n will use Ω as an advice (but it is common for all n and so the algorithm is uniform).

A general sequent in Ω looks like

$$\exists u A(x, y, z, u), \ldots \ \longrightarrow \ \exists v B(x, y, z, v), \ldots$$

To simplify the notation we show just one formula per cedent and we do not show explicit bounds in the existential quantifiers.

The proof β^* will contain n atoms p for bits of x, $n^{O(1)}$ atoms q for bits of y and $t(2^n) \leq 2^{n^{O(1)}}$ atoms r for bits of z. Proof Ω is translated into β^* step by step. If we were constructing a simulation in EF, a sequent of the form as above would be translated into a sequent of the form

$$||A(x, y, z, u)||(p, q, r, u), \ldots \quad \longrightarrow \quad ||B(x, y, z, v)||(p, q, r, v), \ldots$$

where we denote new atoms assigned to bits of u and v ($\leq 2^{n^{O(1)}}$ of them) also u and v for simplicity of the notation. Here u are new atoms that are not extension atoms and are intended to represent bits of a witness to the existential quantifier in the antecedent of the sequent, while v are extension atoms depending possibly on all p, q, r, u. Atoms v are intended to represent bits of a witness to the existential quantifier in the succedent. The fact that v are extension atoms depending on p, q, r, u means that the witness is computed by a circuit from x, y, z, u. The circuit (i.e. the extension axioms) are constructed along with the propositional proof simulating Ω. But as there are exponentially many atoms r already, such a sequent would be exponentially long and could not be produced by a polynomial size circuit.

We overcome this difficulty by systematically introducing new extension atoms for all (sub)formulas that appear in the translation. Hence the sequent gets translated into a sequent of the form

$$w_A, \ldots \longrightarrow w_B, \ldots$$

where w_A and w_B are extension atoms depending on p, q, r, u and p, q, r, v (and hence u too) respectively, and represent the truth values of formulas A and B, respectively.

Having the sequent from Ω this introduction of the extension atoms is exponential in size but very canonical and can be constructed by a polynomial size circuit with an access to Ω. By this phrase we mean that the circuit has size $n^{O(1)}$ and produces the extension atoms and axioms bit by bit (an atom is a letter followed by an index, so the phrase "bit by bit" means that the indices are produced bit by bit).

The whole proof β^* consists of distinct pieces that correspond to sequents in Ω. Each piece has its own canonical assignment, depending only on the sequent but not on how it was derived in Ω, of extension atoms and is constructed by a suitable polynomial size circuit. It remains to show how these pieces are put together to form an EF-proof. That is, how are the inferences in Ω simulated.

We shall consider only the most complicated case, the simulation of a Σ_1^b-LIND inference

$$\frac{\exists u A(t, u) \;\rightarrow\; \exists v A(t+1, v)}{\exists u' A(0, u') \;\rightarrow\; \exists v' A(|w|, v')}$$

(we leave out the free parameters including x, y, z and the quantifier bounds). Assume that the proof β^* contains a derivation of a sequent of the from $w_A \longrightarrow w_B$ representing

$$||A||(t, u) \;\longrightarrow\; ||A||(s, v)$$

where t are new atoms (not extension atoms) representing bits of a witness to the existential quantifier in the antecedent, and s are extension atoms introduced so that they define the number represented by t plus 1 (so their definition just copies a circuit computing the successor function), and v are extension atoms depending on (p, q, r and) t, u representing a witness to the existential quantifier in the succedent (ie. they are computed by a circuit from t, u and from the free parameters).

Take $|w| = 2^{n^{O(1)}}$ copies of this derivation (canonically listed), all written in disjoint copies of atoms t, s, u, v, say t^i, s^i, u^i, v^i for $0 \leq i < 2^{n^{O(1)}}$. The copies copy also the extension axioms. Piece the copies together by postulating that $t^0 = \overline{0}$ (represents 0), that $s^i = t^{i+1}$, and that $v^i = u^{i+1}$. This we can do as atoms t^i and u^i were not extension atoms and so we can add conditions on them to the proof.

This concatenation of the $|w|$ subproofs is again quite canonical and it constitutes a proof of a sequent of the form $w_A \longrightarrow w_B$ corresponding to:

$$||A||(0, u^0) \ \longrightarrow \ ||A||(|w|, v^{|w|}) \ .$$

To finish the description of β^* we only need to derive the (translation of the) antecedent $t(x, y) \leq |z|$ of $(*)$. This is done by stipulating (by extension axioms) that all atoms r are equal to 1 and by using a canonical EF-proof of the valid inequality saying that the term $t(x, y)$ produces from x and y of the lengths n and $n^{O(1)}$, respectively, at most $2^{n^{O(1)}}$ bits.

The EF-proof α of the formula $Correct_\beta$ is easy and uses the splitting of β^* into pieces given by the steps in Ω. It is essentially an EF-proof of the fact that Ω is indeed a proof in $S^1_2 + 1$-Exp of $(*)$.

This concludes the proof of the second part of the theorem.

The third property of the correspondence between iE and V^1_2 stated in the theorem is actually a consequence of the first two (this is a standard argument, cf. [5]). The formalization of the constructions in items 2. and 3. is routine. Note that the formalization starts with Ω and not with an arbitrary V^1_2-proof, i.e. we do not need to formalize the cut-elimination etc. (that would not be possible in S^1_2). This concludes the proof of Theorem 2.1.

q.e.d.

Now we note some corollaries of the theorem. The first one just restates explicitly what has been used in the proof of the theorem (the last sentence in the corollary follows by a general well-known argument using the correspondence between a theory and a proof system).

Corollary 2.2 *Let $\forall x \psi(x)$ be a $\forall \Pi_1^b$-sentence that is provable in $S_2^1 + 1$-Exp, i.e so that S_2^1 proves*

$$|y| \geq t(x) \rightarrow \psi(x) \ .$$

Then the sequence of tautologies $||\psi(x)||^n$, $n \geq 1$, admits polynomial size iEF-proofs.

Moreover, the set of all $\forall \Pi_1^b$-sentences provable in $S_2^1 + 1$-Exp is axiomatized over S_2^1 by the canonical (see [5]) $\forall \Pi_1^b$-sentence expressing the soundness of iEF.

By [4] $S_2^1 + 1$-*Exp* is not $\forall \Pi_1^b$-conservative over S_2^1. Hence Corollary 2.2 immediately yields

Corollary 2.3 *The soundness of iEF is not provable in S_2^1.*

Note that it is not known if S_2^1 proves the soundness of the quantified propositional calculus G.

Theorem 2.1 yields an information about the relative strength of G and iEF.

Corollary 2.4 *iEF p-simulates G.*

Proof :

By [9] the proof system G corresponds to theory U_2^1 and, in particular, the two properties of the correspondence singled out in the introduction are valid for U_2^1 and G. This implies (as U_2^1 is weaker than V_2^1) that V_2^1 proves the soundness of G, and hence iEF polynomially simulates G by the third property stated in Theorem 2.1.

q.e.d.

Proving Corollary 2.4 directly would be rather challenging to a formalization. It is not very difficult to prove directly (via a witnessing style argument) that iEF polynomially simulates G_1. But the simulation of full G, say via Herbrand theorem, would lead to very convoluted formulas (similarly as formulas in Herbrand theorem get complex with the growth of the quantifier complexity).

3 A general definition

In defining the implicit version of a general proof system we return to the original idea of β computing single bits of a proof rather than whole formulas (in fact, a general proof system needs not to operate with formulas at all).

A Q-proof of τ is any string π such that $Q(\pi) = \tau$. Assume that the computation of Q is performed by a deterministic machine running in time n^c; we shall denote it also Q. We will represent the computation of Q on an input of size n by the list of all $t \leq n^c$ instantaneous descriptions of the computation. This list can be represented by an $t \times O(t)$ 0-1 matrix W: the ith row W_i represents the ith instantaneous description.

By increasing t to $O(t)$ we may assume that t is a power of 2 and that W is a $t \times t$ matrix. Let $k := \log(t)$ and let $\beta(i,j)$, $i = (i_1, \ldots, i_k)$ and $i = (j_1, \ldots, j_k)$, be a circuit with $2k$ inputs.

Let $Correct_\beta^Q$ be a canonical propositional formula expressing that:

- The matrix $W_{i,j} := \beta(i,j)$ satisfies all local conditions in order to be a valid computation of Q on an input (encoded in the first row of W).

Note that the size of $Correct_\beta^Q$ is $O(|\beta|)$.

Definition 3.1 *Let P, Q be any proof systems. Define a new proof system $[P, Q]$ as follows. A $[P, Q]$-proof of $\tau \in TAUT$ is a pair (α, β) such that:*

1. *β is a single-output boolean circuit in variables $(i_1, \ldots, i_k, j_1, \ldots, j_k)$, some $k \geq 1$.*
2. *β defines a valid computation W of Q on an input whose output is τ.*
3. *α is a P-proof of the tautology $Correct_\beta^Q$.*

As before we need not to ask for a P-proof of the fact that the output of W is τ. Note also that we could have defined analogously $[P, Q]$-proofs of (possibly exponentially long) formulas τ given implicitly by a circuit; this is taken up in [6].

We would like to define now the implicit version of P to be $[P, P]$. But first we need to verify that this new definition will agree for $P = EF$ with the definition of iEF given in Definition 1.1.

Lemma 3.2 *The proof systems iEF and $[EF, EF]$ polynomially simulate each other, provably in S_2^1.*

Proof :

To conform with the definition of a proof system being a p-time function we will think of EF as being a function computed by the following specific machine. On input π the machine subsequently verifies one step of π after another, checking that the steps are formed from formulas and that they were derived as specified in π. When all these individual checks are fulfilled the machine outputs τ, otherwise it outputs 1.

Let (γ, δ) be an $[EF, EF]$-proof of τ. The computation W defined by δ contains an EF-proof π of τ in it's first row W_1. Hence a circuit β of size computing the steps of π can be readily constructed from δ. However, we need to see that the size of β is $|\delta|^{O(1)}$. In order to achieve this we need to preprocess δ so that π does not contain big formulas. By a general p-simulation (cf. [3]) every EF-proof can be transformed into another one, at most polynomially longer, where all formulas contain at most $|\tau|$ occurrences of atoms. This transformation is very explicit and can be done by a p-time algorithm on the level of circuits describing a proof by a p-time algorithm.

Furthermore, the particular definition of the machine means that the formula $Correct_\beta$ from Definition 1.1 is a simple consequence of $Correct_\delta^{EF}$ (the implication clearly has a p-size EF-derivation). Hence α can be constructed by joining this derivation of $Correct_\beta$ with γ. This shows that $[EF, EF]$ p-simulates iEF.

Now let (α, β) be an iEF-proof of τ, with $|\beta^*| < 2^k$. The individual checks done by the machine computing EF on $\pi := \beta^*$ are parametrized by $i < 2^k$. The ith check can be performed by a fixed p-time algorithm knowing only the formula $\beta^*(i)$ plus the information how it was derived, i.e. knowing only $\beta(i)$. Hence we can construct from β a circuit δ describing this computation W, and $|\delta|$ is $|\beta|^{O(1)}$.

The formula $Correct_\delta^{EF}$ asserts that all local conditions posed on W are met. This is obvious from the construction of δ for the whole of W except for the last part in which the machine collects the results of individual checks and proclaims them all affirmative. In order to prove that this affirmative proclamations are correct we need to know that π was indeed an EF-proof, i.e. we need to prove $Correct_\beta$. However, such a proof is provided by α. So the wanted proof γ of $Correct_\delta^{EF}$ is constructed from α, and $|\gamma|$ is $|\alpha|^{O(1)}$. This shows that $[EF, EF]$ p-simulates iEF.

We leave it to the reader to verify that the simulations can be formalized in S^1_2.

q.e.d.

Definition 3.3 *For any proof system P define* implicit P *to be the proof system* $iP := [P, P]$.

4 Iteration of the construction

We note few simple properties of the bracket operation. The symbols $\leq_p$ and $\equiv_p$ denote the p-simulation and the p-equivalence, respectively. Recall that F, R and R^* denote a Frege system, resolution and tree-like resolution, respectively.

Lemma 4.1 *For all P, $P \leq_p [P, R^*]$.*

Proof :

We will show first that $P \leq_p [P, F]$ and observe at the end that F could be replaced by R^*.

Let $\tau(x_1, \ldots, x_n)$ be a tautology. Circuit β will describe the following trivial, exponential derivation of τ in F (plus the canonical verification that it is an F-proof). For each $a \in \{0,1\}^n$, β^* has a segment where it computes the truth value of $\tau(a)$: this is simply the derivation of subformulas which are true, respectively of the negations of subformulas which are false.

Then it contains 2^{n-1} segments, one for each $(a_2, \ldots, a_n) \in \{0,1\}^{n-1}$ where it derives $\tau(x_1, a_2, \ldots, a_n)$ from $\tau(0, a_2, \ldots, a_n)$ and $\tau(1, a_2, \ldots, a_n)$ (using $x_1 \equiv 0 \vee x_1 \equiv 1$).

Then there are 2^{n-2} segments where all $\tau(x_1, x_2, a_3, \ldots, a_n)$ are derived from $\tau(x_1, 0, a_3, \ldots, a_n)$ and $\tau(x_1, 1, a_3, \ldots, a_n)$, etc. The proof ends with a derivation of τ from $\tau(x_1, \ldots, x_{n-1}, 0)$ and $\tau(x_1, \ldots, x_{n-1}, 1)$.

The correctness of the steps in β^* is trivial to prove assuming one knows that all $\tau(a)$'s have been derived, i.e. are true. But if P proves τ, it proves that "all $\tau(a)$ are true", and hence can prove the formula $Correct_\beta^F$. So $[P, F]$ p-simulates P.

It is easy to see that one can rewrite β^* into a tree-like resolution refutation of (the clauses representing) the formula $\neg\tau$. Hence, in fact, $P \leq_p [P, R^*]$.

q.e.d.

The next lemma shows that it makes no sense to iterate the construction in the place of α.

Lemma 4.2 *For all $P \geq_p EF$, $iP \equiv_p [iP, P]$.*

Proof :

The p-simulation of iP by $[iP, P]$ follows from Lemma 4.1. For the opposite p-simulation consider first the case $P = EF$.

In the proof of the soundness of iEF in V_2^1 we only used the fact that α is an EF-proof in order to know that what α proves is actually true in the model. That is, we only used that the soundness of EF is provable in V_2^1. Hence α could have been an iEF-proof as well. This shows (by part 3 of Theorem 2.1) that $iEF \geq_p [iEF, EF]$.

The case of general $P \geq_p EF$ is proved analogously, using a theory corresponding to iP in place of V_2^1. Such a theory exists by virtue of a general construction of [7]. For example, S_2^1 augmented by a $\forall\Pi_1^b$-sentence asserting the soundness of iP as an extra axiom can be used.

q.e.d.

The restriction to P's p-simulating EF in the lemma is added for technical reasons only. If $P \geq_p EF$ we can appeal to a general construction of a corresponding theory in [7]. But, in fact, such theories exists for many weaker systems like R or F too, cf. [5].

So if we want to iterate the i-construction we should apply it to the second argument in the bracket operation. For the rest of the section we restrict ourselves to $P = EF$.

Definition 4.3 *Put $i_1 EF := iEF$, and for $k \geq 1$ define*

$$i_{k+1}EF \ := \ [EF, i_k EF]$$

One can show analogously to Theorem 2.1 (or by applying Theorem 2.1 to its own formalization in S_2^1) that $i_k EF$ corresponds to $S_2^1 + k\text{-}Exp$ of [4] (or see [5]) and hence to the Σ_1-induction in a k-th order bounded arithmetic. Analogously to Corollary 2.3, $S_2^1 + k\text{-}Exp$ does not prove the soundness of $i_{k+1}EF$. We shall not get into details as we are unable to say anything else sensible about the proof systems besides the next theorem.

Theorem 4.4 *The soundness of each $i_k EF$, $k \geq 1$, is provable in $T_2 + Exp$.*

On the other hand, if a $\forall\Pi_1^b$-sentence $\forall x \psi(x)$ is provable in $T_2 + Exp$ then there is a $k \geq 1$ such that all tautologies $||\psi(x)||^n$, $n \geq 1$, have polynomial size $i_k EF$-proofs.

In the correspondence between $T_2 + Exp$ and $i_k EF$'s the constant k is fixed in proofs of any particular sequence $||\psi(x)||^n$, $n \geq 1$. But we can also allow k unbounded (besides the implicit bound given by the size of the whole proof). In this way we get a proof system that is (presumably) stronger. This is analogous to the situation for G: proofs in T_2 translate into G_k-proofs, fixed $k \geq 1$, while G (unbounded quantifier complexity) corresponds to a stronger theory U_2^1. A formal definition of this very strong proof system might be as follows.

Definition 4.5 *Proof system $i_\infty EF$ is defined as follows. An $i_\infty EF$-proof of $\tau \in TAUT$ is a triple (α, β, w) such that (α, β) is an $i_{|w|}EF$-proof of τ.*

It can be shown that $T_2 + Exp$ does not prove the soundness of $i_\infty EF$. This is an evidence that $i_\infty EF$ may be indeed stronger than any $i_k EF$.

It is easy to see that $i(i_\infty EF) \equiv_p i_\infty EF$ and hence the i-operation does not necessarily always produce a stronger proof system. But we can now start iterating the i_∞-operation and proceed forward. We could have also defined the i_∞-operation not as $|w|$-iteration of the i-operation but as w-iteration (or even 2^w-iteration, etc.) enumerated by a polynomial size circuit (or by a circuit produced by a polynomial size circuit, etc.).

In fact, there does not seem to be a canonical way how to iterate the basic i-operation. This appears analogous to a situation in proof theory of higher order arithmetic and set theory where there is also no canonical way how to iterate consistency statements or even how to represent ordinals.

We conclude by two remarks about the bracket operation for systems below EF, e.g. for F. Consider what would happened if we were to define $[P, F]$ analogously to Definition 1.1, requiring that β outputs whole formulas forming the steps of β^*. To avoid a confusion let us denote this modified bracket operation by $[P, F]^m$.

For example, U_2^1 can be described (its bounded first-order consequences, precisely) as $R_2^1 + 1\text{-}Exp$, where R_2^1 is a subtheory of S_2^1 corresponding to quasipolynomial Frege systems F. But we cannot conclude analogously to Theorem 2.1 that $[F, F]^m$ corresponds to U_2^1. This is because F has no extension atoms and cannot abbreviate a priori exponentially long formulas translating formulas in the starting arithmetical proof, no matter that it is equally canonical as in the case of V_2^1.

The absence of extension atoms in F has another corollary: For any $P \geq_p G_1$ it holds that $[P, F]^m \equiv_p P$. This can be seen as follows. As $P \geq_p G_1$ we can take for a theory T_P corresponding to P (it is unique only up to $\forall\Pi_1^b$-consequences) a theory containing T_2^1. Now assume that (α, β) is an $[P, F]^m$-proof of τ in a model of T_P. The P-proof α is sound in the model and hence β^* is indeed an F-proof of τ. As there are no other atoms in β^* than the atoms of τ, a truth assignment falsifying τ would transfer β^* into a sequence of 0's and 1's which has no first occurrence of 0. That contradicts the minimization principle for Δ_1^b-formulas valid in the model (by T_2^1). Hence T_P proves the soundness of $[P, F]^m$ and so $P \geq_p [P, F]^m$. The opposite simulation $[P, F]^m \geq_p P$ follows by (the proof of) Lemma 4.1 (formulas in β^* there are of polynomial size). In fact, $P \equiv_p [P, F]^m \equiv_p [P, R]^m$.

Acknowledgements: I thank P. Pudlák for discussions over the draft of this paper.

References

[1] S. R. Buss, Bounded Arithmetic. Naples, (1986), Bibliopolis.

[2] S. A. Cook, Feasibly constructive proofs and the propositional calculus, in: *Proc. 7th Annual ACM Symp.on Theory of Computing*, (1975), pp. 83-97. ACM Press.

[3] S. A. Cook and A. R. Reckhow, The relative efficiency of propositional proof systems, *J. Symbolic Logic*, **44(1)**, (1979), pp.36-50.

[4] J. Krajíček, Exponentiation and Second Order Bounded Arithmetic, *Annals of Pure and Applied Logic*, **48(3)**, (1990), pp. 261-276.

[5] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, (1995).

[6] J. Krajíček, Diagonalization in proof complexity, preprint Dec.'03.

[7] J. Krajíček, P. Pudlák, Propositional Proof Systems, the Consistency of First Order Theories and the Complexity of Computations, *J. Symbolic Logic*, **54(3)**, (1989), pp. 1063-1079.

[8] J. Krajíček, P. Pudlák, Quantified propositional calculi and fragments of bounded arithmetic, *Zeitschrift f. Mathematische Logik u. Grundlagen d. Mathematik*, **36**, (1990), pp. 29-46.

[9] J. Krajíček, G. Takeuti, On Bounded $\sum_1^1$-Polynomial Induction, in: *Feasible Mathematics*, eds. S.R. Buss and P.J. Scott, (1990), Birkhauser, pp. 259-280.

Mailing address:
Mathematical Institute
Academy of Sciences
Žitná 25, Prague 1, CZ - 115 67
The Czech Republic
krajicek@math.cas.cz