Unprovability of Herbrand Consistency in Weak Arithmetics

SAEED SALEHI
Institute of Mathematics, Polish Academy of Sciences, P.O.BOX 137, 00-950 Warsaw, and
Turku Center for Computer Sciences, Lemminkaisenkatu 14 A - 4th floor, FIN-20520 Turku
saeed@cs.utu.fi

ABSTRACT. By introducing an appropriate definition of Herbrand Consistency in weak arithmetics, we show Gödel's Second Incompleteness Theorem for Herbrand consistency of theories containing $I\Delta_0$.

1 Introduction

Consider a formula ϕ in the prenex normal form

$$\forall x_1 \exists y_1 \cdots \forall x_m \exists y_m \overline{\theta}(x_1, y_1, \cdots, x_m, y_m)$$

with the Skolem functions $f_1^\theta, \cdots, f_m^\theta$; its Skolemized form by definition is

$$\forall x_1 \cdots \forall x_m \overline{\theta}(x_1, f_1^\theta(x_1), \cdots, x_m, f_m^\theta(x_1, \dots, x_m)).$$

For a sequence of terms $\sigma = \langle t_1, \cdots, t_m \rangle$, the *Skolem instance* $Sk(\theta, \sigma)$ is

$$\overline{\theta}(t_1, f_1^\theta(t_1), \cdots, t_m, f_m^\theta(t_1, \dots, t_m)).$$

Herbrands's Theorem states that a theory T is consistent if and only if every finite set of its Skolem instances is propositionally satisfiable (see (5).) Let Λ be a set of Skolem terms of T (i.e. constructed from the Skolem function symbols of T) *available Skolem instances* of θ in Λ are $Sk(\theta, \sigma)$ for all sequence of terms $\sigma = \langle t_1, \cdots, t_m \rangle \subseteq \Lambda$ such that $\{f_1^\theta(t_1), \cdots, f_m^\theta(t_1, \dots, t_m)\}$ is a subset of Λ too.

Any function, p, whose domain is a set of atomic formulae and its range is $\{0, 1\}$ is called an *evaluation*, if it preserves the equality (for all a, b and atomic formulae φ, $p[a = b] = 1$ implies $p[\varphi(a)] = p[\varphi(b)]$) and satisfies the equality axioms ($p[a = a] = 1$ for all a.) For a set of terms Λ, an *evaluation on* Λ is an evaluation whose domain is the set of all atomic formulae with

Proceedings of the Sixth ESSLLI Student Session
Kristina Striegnitz (editor)
Chapter 23,

constants from Λ (i.e. the variables are substituted by the terms from Λ.) An evaluation p *satisfies* an atomic formula φ if $p[\varphi] = 1$. This definition can be extended to all open (quantifier-less) formulae in a unique way.
Evaluation p on Λ is an T-*evaluation* for a theory T, if it satisfies all the available Skolem instances of T in Λ.
When Λ is the set of all Skolem terms of T, any T-evaluation on Λ determines a Herbrand model of T (see (5).)

Toward formalizing the definition of Herbrand Consistency, we read the above Herbrand's Theorem as:

A theory T is consistent if and only if for every finite set of Skolem terms of T, say Λ, there is an T-evaluation on Λ.

So Herbrand Consistency of a theory T can be defined as:

"for every set of Skolem terms of T, there is an T-evaluation on it."

Herbrand's Theorem is provable in $I\Delta_0 + SupExp$, and it is known that Herbrand consistency is not equivalent to the standard, say Hilbert's, consistency in $I\Delta_0 + Exp$ (see (3), (7).) Unprovability or provability of Herbrand's Consistency for weak arithmetics (i.e. proper fragments of $I\Delta_0 + Exp$) had been an open problem (see (6),(7).) Herbrand Consistency of $I\Delta_0 + Exp$ is unprovable in itself ((3),(7).)

Adamowicz ((1)) has shown the unprovability of Herbrand Consistency of $I\Delta_0 + \Omega_2$ in itself (also in another unpublished paper for $I\Delta_0 + \Omega_1$.)

In this paper we modify the definition of Herbrand Consistency such that its negation gives a real Herbrand proof of contradiction even when Exp is not available, and show unprovability of formalized Herbrand Consistency of $I\Delta_0$ (by the new definition) in itself. So it turns out that $I\Delta_0$ does not prove its own Herbrand Consistency, since the new Herbrand Consistency predicate is implied by the old one.

2 Formalization of Herbrand Consistency in $I\Delta_0$

We take the language of arithmetic $\mathcal{L} = \{0, S, +, ., \leq\}$ in which the operations "S" (successor) "$+$" (addition) and "." (multiplication) are regarded as predicates. For example "$x+y=z$" is a 3-array predicate, and the traditional statements should be re-read in this language by using the predicates $S, +, .$; as an example $\forall x, y, z(x + (y + z) = (x + y) + z)$ can be read as $\forall x, y, z, u, v, w($"$y + z = v$" $\wedge$ "$x + v = w$" $\wedge$ "$x + y = u$" $\rightarrow$ "$u + z = w$"$)$.

So we may need some extra universal quantifiers (and variables) to represent the arithmetical formulae in this language, but for simplicity, and when there is no confusion, we will use the old notation.

All atomic formulae in our language are of the form $x_1 = x_2$, $x_2 = S(x_1)$, $x_1 + x_2 = x_3$, $x_1.x_2 = x_3$ and $x_1 \leq x_2$, where x_1, x_2, x_3 are variables or the constant 0.

Denote the cardinal of a set A by $|A|$; by *terms* we mean terms constructed from the Skolem function symbols of a theory T under consideration.

For a set of terms Λ, there are $2|\Lambda|^3+3|\Lambda|^2$ different atomic formulae with constants from Λ. So there are $2^{2|\Lambda|^3+3|\Lambda|^2}$ different evaluations on Λ. This shows that the above definition has a deficiency in weak arithmetics, from the viewpoint of incompleteness: unprovability of the consistency of T in T is equivalent to having a model of T which contains a proof of contradiction from T. By the above definition, a Herbrand proof of contradiction consists of a set of terms, say Λ, such that there is no T-evaluation on it. If Exp is not available in T, it may happen that all the (few) available evaluations in the model are T-evaluations. This doesn't give a real Herbrand proof in the model! because *not all the evaluations are accessible in the model* (their number $2^{2|\Lambda|^3+3|\Lambda|^2}$ might be too large to exist.) It would be more reasonable if we could find a model with a sufficiently small set of terms in it, such that none of the evaluations on this set (which can be counted in the model) is an T-evaluation. An upper bound for the codes of the evaluations on a set of terms is given below.

We use the Hajek-Pudlak's coding of sets-sequences and terms ((3)) the main properties of this coding are:

$*$ $\mathsf{code}(\langle x_1,\cdots,x_l\rangle), \mathsf{code}(\{x_1,\cdots,x_l\}) \leq (9(1+max\{x_1,\cdots,x_l\})^2)^l$

(i.e. the set $\{x_1,\cdots,x_l\}$ or the sequence $\langle x_1,\cdots,x_l\rangle$ can have a code which is less than or equal to $[9(1+max\{x_1,\cdots,x_l\})^2]^l$)

$*$ $\mathsf{code}(A\cup B), \mathsf{code}(A*B) \leq 64.\mathsf{code}(A).\mathsf{code}(B)$

Code the ordered pair $\langle a,b\rangle$ by $(a+b)^2+b+1$.

Fix the function symbols $f_k^{i,j}$ which is supposed to be the i-th, k-array Skolem function for the j-th axiom of a theory T (so if the j-th axiom is $\exists x\forall y\exists u\exists v A(x,y,u,v)$ then its Skolemized is $\forall y A(f_0^{1,j},y,f_1^{1,j}(y),f_1^{2,j}(y))$.)

Code $f_k^{i,j}$ by $\langle 1,\langle i,\langle j,k\rangle\rangle\rangle$, the symbol ")" by $\langle 2,0\rangle$, "(" by $\langle 2,1\rangle$ and the constant 0 by $\langle 2,2\rangle$.

And fix the function symbols f_l^i which is supposed to be the i-th, l-array function, these symbols are reserved to be Skolem function of a formula θ in the definition of $HCon_T(\theta)$, and code it by $\langle 0,\langle i,l\rangle\rangle$.

Terms are well-bracketing sequence constructed from $\{(,)\}\cup\{f_k^j\}_{j,k}\cup\{f_l\}_l$ (see (3).)

Example Let $i\geq 1$, and define $c_0=0$, $c_{k+1}=f_1^{1,1}(c_k)$ for $k\leq i$. There is a natural number $\mathbf{A}$ such that $\mathsf{code}(\{c_0,\cdots,c_i\})\leq \mathbf{A}^{i^2}$.

Since we have $\mathsf{code}(c_{k+1}) \leq 64^4\mathsf{code}(f_1^{1,1})\mathsf{code}(``(")\mathsf{code}(c_k)\mathsf{code}(``)") \leq 64^4\langle 1,\langle 1,\langle 1,1\rangle\rangle\rangle\langle 2,0\rangle\langle 2,1\rangle\langle 2,1\rangle\mathsf{code}(c_k)$,

let $m=64^4\langle 1,\langle 1,\langle 1,1\rangle\rangle\rangle\langle 2,0\rangle\langle 2,1\rangle\langle 2,1\rangle\mathsf{code}(c_k)$, so we have $\mathsf{code}(c_k)\leq m^k.\mathsf{code}(c_0)$.

Hence $\mathsf{code}(\{c_0,\cdots c_i\}) \leq (9(1+\mathsf{code}(c_i))^2)^i \leq 9^i(2^2(m^i\mathsf{code}(c_0))^2)^i \leq 36^i\langle 2,2\rangle^{2i}m^{2i^2} \leq (36\langle 2,2\rangle m^2)^{i^2}$, we can take $\mathbf{A}=36\langle 2,2\rangle m^2$.

Let Λ be a set of terms with code y, we compute an upper bound for

evaluations on Λ: each evaluation is of the form
$\{\langle y_1 = y_2, p[y_1 = y_2]\rangle \mid y_1, y_2 \in \Lambda\} \cup \{\langle y_1 \leq y_2, p[y_1 \leq y_2]\rangle \mid y_1, y_2 \in \Lambda\} \cup \{\langle y_2 = S(y_1), p[y_2 = S(y_1)]\rangle \mid y_1, y_2 \in \Lambda\} \cup \{\langle y_1.y_2 = y_3, p[y_1.y_2 = y_3]\rangle \mid y_1, y_2, y_3 \in \Lambda\} \cup \{\langle y_1 + y_2 = y_3, p[y_1 + y_2 = y_3]\rangle \mid y_1, y_2, y_3 \in \Lambda\}$;
in which $p[\phi] \in \{0,1\}$ for any atomic formula ϕ with constants from Λ.
Code " = " by $\langle 3,0\rangle$, " $\leq$ " by $\langle 3,1\rangle$, "S" by $\langle 3,2\rangle$, " + " by $\langle 3,3\rangle$, and "." by $\langle 3,4\rangle$.
We code formulae by Polish notation, for example

$\mathsf{code}(x_1 + x_2 = x_3) = \mathsf{code}(+(x_1.x_2.x_3)) =$

$\mathsf{code}(\langle\langle 3,3\rangle, \langle 2,0\rangle, \mathsf{code}(x_1), \mathsf{code}(x_2), \mathsf{code}(x_3), \langle 2,1\rangle\rangle)$.

There is a natural number $\mathbf{a}$ such that for any $k \in \{0,1\}$

$\mathsf{code}(\langle y_1 = y_2, k\rangle) \leq 2 + (1 + \mathbf{a}y_1y_2)^2$,

$\mathsf{code}(\langle y_1 \leq y_2, k\rangle) \leq 2 + (1 + \mathbf{a}y_1y_2)^2$,

$\mathsf{code}(\langle y_2 = S(y_1), k\rangle) \leq 2 + (1 + \mathbf{a}y_1y_2)^2$,

$\mathsf{code}(\langle y_1 + y_2 = y_3, k\rangle) \leq 2 + (1 + \mathbf{a}y_1y_2y_3)^2$, and

$\mathsf{code}(\langle y_1.y_2 = y_3, k\rangle) \leq 2 + (1 + \mathbf{a}y_1y_2y_3)^2$.

So $\mathsf{code}(\langle \phi, k\rangle) \leq 2 + (1 + \mathbf{a}y^3)^2$ for all $k \in \{0,1\}$ and atomic ϕ with constants from Λ, with $\mathsf{code}(\Lambda) = y$. Hence $\mathsf{code}(p) \leq (9.(3 + (1 + \mathbf{a}y^3)^2)^2)^{2|y|^3+3|y|^2} \leq (81(1 + \mathbf{a}y^3)^4)^{2|y|^3+3|y|^2}$, (we identify $|\Lambda|$ with $|y|$) for all evaluation p on Λ. Call a set of terms Λ with $\mathsf{code}(\Lambda) = y$, *admissible* if $F(y) = (81(1 + \mathbf{a}y^3)^4)^{2|y|^3+3|y|^2}$ exists.
We modify the definition of Herbrand Consistency of a theory T as: " for every admissible set of Skolem terms of T, there is an T-evaluation on it". This is formalized below.

By "terms" we mean terms constructed from the Skolem function symbols $\{f_k^{i,j}\}_{i,j,k} \cup \{f_l^i\}_{i,l}$ introduced above, the bounded formula $\mathsf{Terms}(y)$ means "y is a set of terms constructed from those symbols".

There are bounded formulae $\mathsf{eva}(x)$ and $\mathsf{eval}(x,y)$ which represent "x is an evaluation" and "y is a set of terms and x is an evaluation on y".

For atomic formula ϕ, $p[\phi] = 1$ is a bounded formula, for more complex ϕ the statement $p[\phi] = 1$ can be written by a Π_1-formula:

let the bounded formula $\mathsf{Sat}(p, \phi, s)$ be

"$\mathsf{eva}(p)$& s is a sequence of pairs $\langle a_i, b_i\rangle$, such that:

1) each a_i is (the code of) a formula and each b_i is 0 or 1,

2) for $k = \mathsf{length}(s)$, $a_k = \phi$ and $b_k = 1$,

3) each a_i is either of the form

3.1) $a_i = a_j \wedge a_k$ for some $j, k < i$ and $b_i = b_j.b_k$,

or 3.2) $a_i = a_j \vee a_k$ for some $j, k < i$ and $b_i = b_j + b_k - b_j.b_k$,

or 3.3) $a_i = a_j \rightarrow a_k$ for some $i, j < k$ and $b_i = 1 + b_j.b_k - b_j$,

or 3.4) $a_i = \neg a_j$ for some $j < i$ and $b_i = 1 - b_j$,

or 3.5) a_i is atomic and $b_i = p[a_i]$. "

Let $S(\theta)$ be the number of subformulae of the formula θ. For the above sequence s we have $\mathsf{code}(s) \leq (9(1 + \mathsf{code}(\langle \phi, 1\rangle))^2)^{S(\phi)}$

$\leq (9(1 + 2 + (\phi + 1)^2)^2)^{S(\phi)} \leq (81(1 + \phi)^4)^{S(\phi)}$.

So we can write $p[\phi]=1$ as: $\forall z\Big(z\geq(81(1+\phi)^4)^{S(\phi)}\rightarrow\exists s\leq zSat(p,\phi,s)\Big)$.

Let $|\theta|$ be the number of existential quantifiers in the prenex normal form of θ (we can assume it has the form $\theta=\forall x_1\exists y_1\cdots\forall x_m\exists y_m\overline{\theta}(x_1,y_1,\cdots,x_m,y_m)$, so $|\theta|=m$ in this case.)
For a formula θ fix its Skolem functions as $f_1^\theta,\cdots,f_\alpha^\theta$ where $\alpha=|\theta|$. Write $\sigma=\langle t_1,\cdots,t_\alpha\rangle\subseteq\Lambda$ for a set of terms Λ such that $\{f_1^\theta(t_1),\cdots,f_\alpha^\theta(t_1,\ldots,t_\alpha)\}$ is a subset of Λ too.

We have $\mathsf{code}(Sk(\theta,\sigma))\leq\mathsf{code}(\theta*\sigma*(f_1^\theta(t_1),\cdots,f_\alpha^\theta(t_1,\ldots,t_\alpha)))$.

On the other hand
$\mathsf{code}((f_1^\theta(t_1),\cdots,f_\alpha^\theta(t_1,\ldots,t_\alpha))))\leq 18^\alpha\mathsf{code}(f_\alpha^\theta(t_1,\ldots,t_\alpha)))^{2\alpha}$, and also $\mathsf{code}(f_\alpha^\theta(t_1,\ldots,t_\alpha)))\leq 64^{3+\alpha}\mathsf{code}(f_\alpha^\theta)\mathsf{code}(\text{“(”})\mathsf{code}(\text{“)”})\mathsf{code}(t_1)\ldots\mathsf{code}(t_\alpha)$.

So with $\mathsf{code}(\Lambda)=y$ we have

$\mathsf{code}(Sk(\theta,\sigma))\leq 64^3\theta(18.55^2)^\alpha 64^{2\alpha(3+\alpha)}\mathsf{code}(f_\alpha^\theta)^{2\alpha}.y^{2\alpha^2+\alpha}$.

Let $G(\theta,y)=[81(1+64^3.(18.55^2)^{|\theta|}.64^{2|\theta|(3+|\theta|)}.\theta.\mathsf{code}(f_{|\theta|}^\theta)^{2|\theta|}.y^{2|\theta|^2+|\theta|})^4]^{S(\theta)}$.

Noting that “$u=Sk(\theta,y)$” is a bounded formula, we can write “p is an θ-evaluation on y” as:

$\mathsf{Terms}(y)\wedge\mathsf{eval}(p,y)\wedge\forall z[z\geq G(\theta,y)\rightarrow\forall u\leq z\forall\sigma\leq y\{\sigma=(t_1,\cdots,t_{|\theta|})\subseteq y\wedge\{f_1^\theta(t_1),\cdots,f_{|\theta|}^\theta(t_1,\ldots,t_\alpha))\}\subseteq y\wedge\text{“}u=Sk(\theta,\sigma)\text{”}\rightarrow\exists s\leq z\mathsf{Sat}(p,u,s)\}]$.
Denote its bounded counterpart by $\mathsf{SatAvail}(p,y,\theta,z)$, that is:

$\mathsf{SatAvail}(p,y,\theta,z)=\mathsf{Terms}(y)\wedge\mathsf{eval}(p,y)\wedge\forall u\leq z\forall\sigma\leq y\{\sigma=(t_1,\cdots,t_{|\theta|})\subseteq y\wedge\{f_1^\theta(t_1),\cdots,f_{|\theta|}^\theta(t_1,\ldots,t_\alpha))\}\subseteq y\wedge\text{“}u=Sk(\theta,\sigma)\text{”}\rightarrow\exists s\leq z\mathsf{Sat}(p,u,s)\}]$.
For a finite theory $\{T_1,\cdots,T_n\}$, define the predicate $HCon_T(x)$, as:
$\forall z\Big(\forall y\leq z\ [\ \mathsf{Terms}(y)\ \wedge\ z\geq F(y)\ \wedge\ \bigwedge_{1\leq j\leq n}z\geq G(T_j,y)\ \wedge\ z\geq G(x,y)\rightarrow$
$\exists p\leq z\exists s\leq z\{\mathsf{eval}(p,y)\wedge\bigwedge_{1\leq j\leq n}\mathsf{SatAvail}(p,T_j,y,s)\wedge\mathsf{SatAvail}(p,x,y,s)\}]\Big)$.

We note that the bounds $G(T_j,y)$ and for a standard x the bound $G(x,y)$ for z, are polynomial with respect to y, so for large enough, also for non-standard ys, they are less than the bound $F(y)$.

The cut log^2 is defined (informally) by: $x\in log^2\iff 2^{2^x}$ *exists*. A formal definition is given in the next section.

The predicate $HCon_T^*(x)$ is obtained from $HCon_T(x)$ by restricting the (only unbounded) universal quantifier to log^2:

$\forall z\in log^2\Big(\forall y\leq z\ [\ \mathsf{Terms}(y)\ \wedge\ z\geq F(y)\ \wedge\ \bigwedge_{1\leq j\leq n}z\geq G(T_j,y)\ \wedge\ z\geq G(x,y)\ \rightarrow\ \exists p\leq z\exists s\leq z\{\mathsf{eval}(p,y)\wedge\bigwedge_{1\leq j\leq n}\mathsf{SatAvail}(p,T_j,y,s)\ \wedge\ \mathsf{SatAvail}(p,x,y,s)\}]\Big)$.

Proposition 2.1. *The above formulae $HCon_T(\phi)$ and $HCon_T^*(\phi)$ binumerate “Herbrand Consistency of T with ϕ” in $\mathbf{N}$:*

$\mathbf{N}\models HCon_T(\theta)$ *iff* $\mathbf{N}\models HCon_T^*(\theta)$ *iff “$\{\phi\}\cup T$ is Herbrand consistent.”*

Herbrand Consistency of T, $HCon(T)$, is $HCon_T(\text{“}0=0\text{”})$.

For a moment assume we have proved the following proposition:

Proposition 2.2. *There is a finite set of $I\Delta_0$-derivable sentences, say B, such that for every bounded formula $\theta(x)$ with x as the only free variable, and for any finite theory α (in the language of arithmetic) whose axioms contain the set B,*

$$I\Delta_0 \vdash HCon(\alpha) \wedge \exists x \in log^2\theta(x) \rightarrow HCon^*_\alpha(\text{“}\exists x \in log^2\theta(x)\text{”})$$

Now we can prove our main theorem:

Theorem 2.3. *Take B as the previous proposition, and let D be the union of B and a finite fragment of $I\Delta_0$ containing PA^- such that the last proposition is provable in D, then for any finite consistent theory (in the language of arithmetic) whose axioms contain the set D, we have $\alpha \nvdash HCon(\alpha)$.*

Proof. Let τ be the fixed point of $HCon^*_\alpha(\neg\tau) \equiv \tau$ (it is available in PA^-, i.e. $PA^- \vdash HCon^*_\alpha(\neg\tau) \equiv \tau$, see (4).)

The theory $\alpha + \neg\tau$ is consistent, since otherwise, by proposition 2.1, we would have $\mathbf{N} \models \neg HCon^*_\alpha(\neg\tau)$ and so by the fact that PA^- is Σ_1-complete ((4)) we would get $PA^- \vdash \neg HCon^*_\alpha(\neg\tau)$, hence $\alpha \vdash \neg\tau$, then α would be inconsistent.

Write $\neg\tau \equiv \exists x \in log^2\theta(x)$ for a bounded θ, then

$\alpha + \neg\tau + HCon(\alpha) \vdash HCon(\alpha) \wedge \exists x \in log^2\theta(x)$,

so by proposition 2.2 we get $\alpha + \neg\tau + HCon(\alpha) \vdash HCon^*_\alpha(\text{“}\exists x \in log^2\theta(x)\text{”})$,

and then $\alpha + \neg\tau + HCon(\alpha) \vdash HCon^*_\alpha(\neg\tau)$, or $\alpha + \neg\tau + HCon(\alpha) \vdash \tau$.

So $\alpha \vdash HCon(\alpha) \rightarrow \tau$, and this shows that $\alpha \nvdash HCon(\alpha)$. $\triangle$

3 A Herbrand Σ_1-Completeness Theorem in $I\Delta_0$

This section is devoted to prove proposition 2.2.

Godel's original second incompleteness theorem states unprovability of (formalized) consistency of T in T, for strong enough theories T. Being strong enough means being able to code sets-sequences, terms and some other logical (syntaical) concepts, like provability and prove their properties.

Of those properties are:

1. $T \vdash Pr_T(\varphi) \wedge Pr_T(\varphi \rightarrow \psi) \rightarrow Pr_T(\psi)$
2. $T \vdash Pr_T(\varphi) \rightarrow Pr_T(Pr_T(\varphi))$

Usually the property 2 is proved by use of formalized Σ_1-completeness theorem:

$T \vdash \varphi \rightarrow Pr_T(\varphi)$ for any Σ_1-formula φ.

So how can one show Godel's second incompleteness theorem for weak arithmetics, which are not that strong to prove those properties?

One may have two options here:

1) try to find a model of T which does not satisfy $Con(T)$,

2) try to show some weak forms of Σ_1-completeness in T, which can prove $T \nvdash Con(T)$ (by a similar argument of our main theorem's proof.)

The first method is applied in (2) to show $Q \nvdash Con(Q)$ for Robinson's arithmetic Q.

There is no hope to use this way for more complex theories like $I\Delta_0$ (and its super-fragments) since there is no recursive non-standard model for them (see (3).)

So the difficulty rises when one seeks for a kind of formalized Σ_1-completeness theorem which can be proved in the (weak) theory and at the same time is powerful enough to show unprovabolity of the theory's consistency in itself.

A weak form of Σ_1-incompleteness theorem can be like:

$T \vdash Con(T) \wedge \exists x \theta(x) \rightarrow Con_T(\exists x \theta(x))$ for Δ_0-formulae $\theta(x)$ (cf (1).)

Our proposition 2.2 is a form of weak Σ_1-incompleteness theorem, in which the witness x for $\theta(x)$ is small (restricted to log^2) and the second consistency predicate is rather weak (that is $HCon^*_T$ instead of $HCon_T$.)

Take A be the axiom system:

$A1.$ $\forall x \exists y$ "$y = S(x)$"

$A2.$ $\forall x, y, z($"$y = S(x)$" $\wedge$ "$z = S(x)$" $\rightarrow y = z)$

$A3.$ $\forall x$ $(x \leq x)$

$A4.$ $\forall x, y, z$ $(x \leq y \wedge y \leq z \rightarrow x \leq z)$

$A5.$ $\forall x$ $(x \leq 0 \rightarrow x = 0)$

$A6.$ $\forall x, y, z$ ("$y = S(z)$" $\wedge\, x \leq y \rightarrow x \leq z \vee x = y)$

$A7.$ $\forall x, y($"$y = S(x)$" $\rightarrow x \leq y)$

$A8.$ $\forall x$ "$x + 0 = x$"

$A9.$ $\forall x, y, z, u, v$ ("$z = S(y)$" $\wedge$ "$x + y = u$" $\wedge$ "$v = S(u)$" $\rightarrow$ "$x + z = v$")

$A10.$ $\forall x$ "$x.0 = 0$"

$A11.$ $\forall x, y, z, u, v$ ("$z = S(y)$" $\wedge$ "$x.y = u$" $\wedge$ "$u + x = v$" $\rightarrow$ "$x.z = v$")

$A12.$ $\forall x, y$ ("$y = S(x)$" $\rightarrow \neg y \leq x)$

Fix the terms $c_0 = 0$, $c_{j+1} = f_1^{1,1}(c_j)$.

The term c_i is represented as the i-th numeral in every A-evaluation p: $p[c_0 = 0] = 1$ and $p[c_{j+1} = S(c_j)] = 1$.

Lemma 3.1. *($I\Delta_0$) Suppose for an i, we have $\{c_0, \cdots, c_i\} \subseteq \Lambda$ for a set of terms Λ, and p is an A-evaluation on Λ, then*
1) If $p[a \leq c_i] = 1$ for an $a \in \Lambda$, then there is an $j \leq i$ such that $p[a = c_j] = 1$.
2) If γ is an open formula and $\gamma(x_1, \cdots, x_m)$ holds for $x_1 \cdots x_m \leq i$, then $p[\gamma(c_{x_1}, \cdots, c_{x_m})] = 1$.

Proof. 1) by induction on j, one can prove that if $p[a \leq c_j] = 1$ then $p[a = c_k] = 1$ for a $k \leq j$: for $j = 0$ use $A5$, and for $j + 1$ use $A6$.

2) The assertion can be proved for the atomic or negated atomic formulae. For $x_1 \leq x_2$ use induction on x_2, for $x_2 = 0$ by $A3$ and for $x_2 + 1$ by

$A3$, $A4$ and $A7$. Similarly for $x_1 + x_2 = x_3$ and $x_1.x_2 = x_3$ use induction on x_2 and $A8$, $A9$, $A10$ and $A11$. For $\neg x_1 = x_2$: if $\neg x_1 = x_2$ then either $x_1 + 1 \leq x_2$ or $x_2 + 1 \leq x_1$, e.g. for $x_1 + 1 \leq x_2$ we have $p[c_{x_1+1} \leq c_{x_2}] = 1$, now use $A12$. For $\neg S(x_1) = x_2$ use $A2$, and the cases $\neg x_1 + x_2 = x_3$ and $\neg x_1.x_2 = x_3$ can be derived from the previous cases. For $\neg x_1 \leq x_2$: if $\neg x_1 \leq x_2$ then $x_2 + 1 \leq x_1$ so $p[c_{x_2+1} \leq c_{x_1}] = 1$, now use $A4$ and $A12$.

The induction cases for $\wedge, \vee, \rightarrow$ are straightforward. (Note we have assumed that the formula θ is in normal form: the negation appears only in front of atomic formulas.) $\triangle$

Recall Godel's beta function:
$\beta(a, b, i) = r$ if $a = (q + 1)[(i + 1)b + 1] + r \ \wedge \ r \leq (i + 1)b$ for some q (cf (4).) Define the ordered pairs by $\langle a, b\rangle = a + \frac{1}{2}(a + b + 1)(a + b)$.

Let $\Psi(z, i) = \forall x \leq z \forall y \leq z \forall j < i\{\langle x, y\rangle = z \rightarrow x \geq (i + 1)y + 1 \wedge \beta(x, y, 0) = 2 \wedge \beta(x, y, j + 1) = (\beta(x, y, j))^2\}$.

The formula $\Psi(z, i)$ states that z is a (β)-code of a sequence whose length is at least $i + 1$, and its first term is 2 and every term is the square of its preceding term. So such a sequence looks like: $\langle 2, 2^2, 2^{2^2}, \cdots, 2^{2^i}, \ldots\rangle$.

We can define the cut log^2 as: $x \in log^2 \iff \exists z \Psi(z, x)$.

Denote the open part of Ψ by $\overline{\Psi}$, so $\Psi(z, x) = \forall \mathbf{u} \overline{\Psi}(z, x, \mathbf{u})$, in which $\mathbf{u} = (u_1, \cdots, u_k)$ for a natural k.

To get the B asserted in the proposition, we add the following axioms to A:

$A13$. $\Psi(33, 0)$

$A14$. $\forall x \forall i \exists y (\Psi(x, i) \rightarrow \Psi(y, i + 1))$

The axiom $A14$ is in fact the $I\Delta_0$-derivable statement $i \in log^2 \rightarrow i+1 \in log^2$. To be more precise we (can) write the axiom $A14$ in the prenex normal form:

$A14$. $\forall x \forall i \exists y \forall \mathbf{u} \forall \mathbf{v} (\overline{\Psi}(x, i, \mathbf{u}) \rightarrow \overline{\Psi}(y, i + 1, \mathbf{v}))$.

Fix the terms $z_0 = c_{33}$, $z_{j+1} = f_2^{1,14}(z_j, c_j)$.

The term z_i is represented as a (β)-code of the sequence $\langle 2, 2^2, \cdots, 2^{2^i}\rangle$ in any B-evaluation (note that $33 = \langle 5, 2\rangle =$ a β-code for $\langle 2\rangle$.)

Lemma 3.2. *($I\Delta_0$) Suppose for $i \geq 33$, $\{c_0, \cdots, c_i, z_0, \cdots, z_i\} \subseteq \Lambda$, then for any B-evaluation p on Λ, p satisfies all the available Skolem instances of $\Psi(z_i, c_i)$.*

Proof. By induction on $j \leq i$ one can show that any such p satisfies all the available Skolem instances of $\Psi(z_j, c_j)$. $\triangle$

Now we are close to the proof of the proposition, let α be a theory whose axioms contain the set B, and take a model $M \models I\Delta_0$ such that $M \models HCon(\alpha)$ and $M \models i \in log^2 \wedge \theta(i)$ for an $i \in M$. Take a set of terms Λ with $\mathsf{code}(\Lambda) = y$ such that $F(y)$ exists and is in $log^2(M)$ (we can assume i and y are non-standard) then we find an admissible set of terms Λ', so by the assumption $HCon(\alpha)$ there is an α-evaluation on Λ' which induces

an $(\alpha \cup \{\exists x \in log^2\theta(x)\})$-evaluation on Λ. This shows $M \models HCon^*_\alpha(\exists x \in log^2\theta(x))$.
Write $\theta(x) = \forall x_1 \leq \alpha_1 \exists y_1 \leq \beta_1 \cdots \forall x_m \leq \alpha_m \exists y_m \leq \beta_m \overline{\theta}(x, x_1, y_1, \cdots, x_m, y_m)$. There are (partial) functions on M, $g_1, \cdots, g_m$ (we may assume, $g_j : [0, i]^j \to M$) such that for all $a_1, \cdots, a_m \in M$

$M \models a_1 \leq \alpha'_1 \to [g_1(a_1) \leq \beta'_1 \wedge \cdots [a_m \leq \alpha'_m \to [g_m(a_1, \ldots, a_m) \leq \beta'_m \wedge \overline{\theta}(i, a_1, g_1(a_1), \cdots, g_m(a_1, \ldots, a_m))]] \ldots],$

in which $(\alpha'_j, \beta'_j;\ j \leq m)$ is the image of $(\alpha_j, \beta_j;\ j \leq m)$ under the substitution $\{x \mapsto i, x_j \mapsto a_j, y_j \mapsto g_j(a_1, \cdots a_j);\ j \leq m\}$.

Consider the formula
$\exists x \in log^2\theta(x) \equiv$
$\exists x \exists z \forall x_1 \leq \alpha_1 \exists y_1 \leq \beta_1 \cdots \forall x_m \leq \alpha_m \exists y_m \leq \beta_m \forall \mathbf{u}\{\overline{\Psi}(z, x, \mathbf{u}) \wedge$
$\wedge \overline{\theta}(x, x_1, y_1, \cdots, x_m, y_m)\}$. Its Skolemized form is:

$\forall x_1 \cdots \forall x_m \forall \mathbf{u}\{\overline{\Psi}(f^2_0, f^1_0, \mathbf{u}) \wedge x_{\leq}\alpha''_1 \to [f^1_1(x_1) \leq \beta''_1 \wedge \cdots [x_m \leq \alpha''_m \to [f^1_m(x_1, \ldots, x_m) \leq \beta''_m \wedge \overline{\theta}(f^1_0, x_1, f^1_1(x_1), \cdots, x_m, f^1_m(x_1, \ldots, x_m))]] \cdots]\},$

in which $(\alpha''_j, \beta''_j;\ j \leq m)$ is the image of $(\alpha_j, \beta_j;\ j \leq m)$ under the substitution $\{x \mapsto f^1_0, y_j \mapsto f^1_j(x_1, \cdots x_j);\ j \leq m\}$.

Define the operation $\Re$ on terms by:

- $f^1_0 \mapsto c_i$
- $f^2_0 \mapsto z_i$
- $f^1_1(c_j) \mapsto c_{g_1(j)}$

$\vdots$

- $f^1_m(c_{j_1}, \cdots, c_{j_m}) \mapsto c_{g_m(j_1, \cdots, j_m)}$

That is the term f^1_0 is mapped (under $\Re$) to c_i, and f^2_0 is mapped to z_i and for any $1 \leq t \leq m$ the term $f^1_t(c_{j_1}, \cdots, c_{j_t})$ is mapped to $c_{g_t(j_1, \cdots, j_t)}$.

By an argument similar to the example in the previous section, it can be shown that there is a natural $\mathbf{K}$ such that $\mathsf{code}(c_j), \mathsf{code}(z_j) \leq \mathbf{K}^j$ for any $j \geq 1$, and $\mathsf{code}(\{c_0, \cdots, c_i, z_0, \cdots, z_i\}) \leq \mathbf{K}^{i^2}$ for any $i \geq 1$.
For any term t, $\mathsf{code}(\Re(t)) \leq \mathsf{code}(t * (z_i)^{|t|} * (c_i)^{|t|}) \leq$

$\leq 64^3.t.36^{3|t|}.\mathsf{code}(z_i)^{|t|}.\mathsf{code}(c_i)^{|t|} \leq 64^3.t.36^{3t}.\mathbf{K}^{2it}$, so $max\{\mathsf{code}(\Re(t)) | t \in \Lambda\} \leq 64^3.y.36^{3y}.\mathbf{K}^{2i^2y}$, hence $\mathsf{code}(\Re(y)) \leq 36^{|y|}.[64^3.t.36^{3t}.\mathbf{K}^{2it}]^{|y|}$.

Let $\Lambda' = \Re(\Lambda) \cup \{c_0, \cdots, c_i, z_0, \cdots, z_i\}$.

So $\mathsf{code}(\Lambda') = y' \leq 64.\mathbf{K}^{2i^2}.36^{|y|}.64^{3|y|}.y^{|y|}.36^{3y|y|}.\mathbf{K}^{2i|y|}$.

We show that $F(y')$ exists. Note that $y \in log^2$ because $y < F(y) \in log^2$.

Assuming that i, y are non-standard we can write: $F(y') \leq (y')^{4|y'|^4} = (y')^{4(2i+|y|)^4} \leq (y')^{4y^5}.(y')^{4(2i)^5}$, and this is less than $(2^{2^i})^{14}$ if $y \leq i$, and is less than $(2^{2^y})^{14}$ if $i \leq y$. So Λ' is admissible.

Hence by the assumption $HCon(\alpha)$ there is an α-evaluation q on Λ'. Define the evaluation p on Λ by

$p[\varphi(a_1, \cdots, a_l)] = q[\varphi(\Re(a_1), \cdots, \Re(a_l))]$ for any atomic φ.

It can be shown that the above equality holds for open formulae φ as well. We show that p satisfies all the available Skolem instances of $\{\exists x \in log^2\theta(x)\} \cup \alpha$ in Λ:

1) p is an α-evaluation, since q is so and the operation $\Re$ has nothing to do with the Skolem functions of α: $p[\phi(t_1, f_1^{1,j}(t_1), \cdots, t_k, f_k^{1,j}(t_1, \ldots, t_k))] =$
$q[\phi(\Re(t_1), \Re(f_1^{1,j}(t_1)), \cdots, \Re(t_k), \Re(f_k^{1,j}(t_1, \ldots, t_k)))] =$
$q[\phi(\Re(t_1), f_1^{1,j}(\Re(t_1)), \cdots, \Re(t_k), f_k^{1,j}(\Re(t_1, \ldots, t_k)))] = 1.$
2) p satisfies all the available Skoelm instances of $\exists x \in log^2\theta(x)$ in Λ:
2.1) $p[\overline{\Psi}(f_0^2, f_0^1, t_1, \cdots, t_{24})] = q[\overline{\Psi}(\Re(f_0^2), \Re(f_0^1), \Re(t_1), \cdots, \Re(t_{24}))] =$
$q[\overline{\Psi}(z_i, c_i, \Re(t_1), \cdots, \Re(t_{24}))] = 1$
since by lemma 3.2, q satisfies all the available Skolem instances of $\Psi(z_i, c_i)$ then the latter equality holds.
2.2) by lemma 3.1 for any term t and any $k \leq i$, if $p[t \leq c_k] = 1$ then $p[t = c_j] = 1$ for some $j \leq k$. So for evaluating $\theta(x)$ it is enough to consider Skolem instances like $\bar{\theta}(f_0^1, c_{j_1}, f_1^1(c_{j_1}), \cdots, c_{j_m}, f_m^1(c_{j_1}, \ldots, c_{j_m}))$:
$p[\bar{\theta}(f_0^1, c_{j_1}, f_1^1(c_{j_1}), \cdots, c_{j_m}, f_m^1(c_{j_1}, \ldots, c_{j_m}))] =$
$q[\bar{\theta}(\Re(f_0^1), \Re(c_{j_1}), \Re(f_1^1(c_{j_1})), \cdots, \Re(c_{j_m}), \Re(f_m^1(c_{j_1}, \ldots, c_{j_m})))] =$
$q[\bar{\theta}(c_i, c_{j_1}, c_{g_1(j_1)}, \cdots, c_{j_m}, c_{g_m(j_1, \ldots, j_m)})] = 1$
the latter equality holds by $M \models \bar{\theta}(i, j_1, g_1(j_1), \cdots, j_m, g_m(j_1, \ldots, j_m))$ and lemma 3.1.

This completes the proof of the proposition.

Acknowledgement I would like to thank professor Zofia Adamowicz for reading the draft of the paper and for her fruitful criticisms.

References

[1] Adamowicz Z. & Zbierski P. "On Herbrand Type Consistency in Weak Theories", to appear in *Archive for Mathematical Logic.*

[2] Bezboruah A. & Shepherdson J.C. "Gödel's Second Incompleteness Theorem For Q", in *The Journal of Symbolic Logic.* (**41**) 1976, pp. 503-512

[3] Hájek P. & Pudlák P. *Metamathematics of First Order Arithmetic*, Springer-Verlag 1991.

[4] Kay R. *Models of Peano arithmetic*, Oxford Logic Guides 15,. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1991

[5] Nerode A. & Shore R.A. *Logic for Applications*, Springer-Verlag 1993.

[6] Paris J. & Wilkie A. "Δ_0-sets and inducton" Proceedings of the Jadwisin Logic Conference, Open days in Model Theory and Set Theory, Poland, Leeds University Press, 1981, pp. 237-248.

[7] Pudlák P. "Cuts, Consistency Statements and interpretation" in *The Journal of Symbolic Logic*, (**50**) 1985, pp. 423-442.