Notre Dame Journal of Formal Logic
Volume 35, Number 3, Summer 1994

A Simple Proof of Arithmetical Completeness for Π_1-conservativity Logic

GIORGI JAPARIDZE

Abstract Hájek and Montagna proved that the modal propositional logic *ILM* is the logic of Π_1-conservativity over sound theories containing $I\Sigma_1$ (*PA* with induction restricted to Σ_1 formulas). I give a simpler proof of the same fact.

1 Introduction By a "theory" we mean an effectively axiomatized theory whose language contains that of *PA* (arithmetic).

We say that a theory T_2 is Π_1-*conservative* over a theory T_1 if T_1 proves every Π_1-theorem of T_2. And T_2 is *interpretable* in T_1 if, intuitively, the language of T_2 can be translated into the language of T_1 in such a way that T_1 proves the translation of every theorem of T_2.

We say that a theory is *essentially reflexive* if for any formula α it proves $Pr_{PC}(\ulcorner\alpha\urcorner) \to \alpha$, where $\ulcorner\alpha\urcorner$ is the code (Gödel number) of α and $Pr_{PC}(x)$ is the standard formalization of "x is the code of a formula provable in the classical predicate calculus."

It is known that *PA* is essentially reflexive, but no finitely axiomatizable reasonable theory, including $I\Sigma_1$ (*PA* with induction restricted to Σ_1-formulas), can be such. Indeed, suppose T is a sufficiently strong finitely axiomatized theory. Let then Ax be the conjunction of the universal quantifier closures of its axioms. If T is essentially reflexive, then $T \vdash Pr_{PC}(\ulcorner\neg Ax\urcorner) \to \neg Ax$, whence $T \vdash \neg Pr_{PC}(\ulcorner\neg Ax\urcorner)$, which means that T proves its own consistency and hence by Gödel's Second Incompleteness Theorem T is inconsistent.

According to a nice fact known as *Orey-Hájek characterization*, if given theories are essentially reflexive, one is interpretable in another if and only if one is Π_1-conservative over the other; moreover, this fact is provable in *PA*, so we can say that interpretability and Π_1-conservativity relations between essentially reflexive theories are "the same." However, this is not true for finitely axiomatized theories like $I\Sigma_1$.

De Jongh and Veltman [5] introduced the propositional modal logic *ILM*, whose language contains two modal operators: $\Box$ (unary) and $\rhd$ (binary). Berarducci [1] and

Received May 27, 1994; revised October 18, 1994

Shavrukov [7], independently, proved that *ILM* is the logic of interpretability over *PA*, that is, *ILM* yields exactly the schemata of *PA*-provable formulas, when $\Box A$ is understood as a formalization of "A is *PA*-provable" and $A \rhd B$ as a formalization of "*PA*+B is interpretable in *PA*+A." By the Orey-Hájek characterization, this result immediately implies that *ILM* is the logic of Π_1-conservativity over *PA* as well. However, the question whether *ILM* is the logic of Π_1-conservativity over $I\Sigma_1$ (whose logic of interpretability was in Visser [10] shown to be different from *ILM*) remained open until Hájek and Montagna [6] found a positive answer.

In this paper I present an alternative proof of completeness of *ILM* as the logic of Π_1-conservativity over $I\Sigma_1$ and its sound extensions; this proof is more direct (as it appeals only to the most elementary facts about Π_1-sentences and is based directly on the natural semantics for *ILM*—Veltman models) and therefore considerably simpler than that of Hájek and Montagna; since, in view of the Orey-Hájek characterization, this result immediately implies completeness of *ILM* as the logic of interpretability over *PA*, this is at the same time a new proof of the above-mentioned Berarducci-Shavrukov theorem, which seems the simplest among those known so far.

2 Modal Logic Preliminaries

ILM is given as the classical propositional logic plus the rule of necessitation $\vdash A \Rightarrow \ \vdash \Box A$ and the following axiom schemata ($\Diamond = \neg\Box\neg$):

$$\Box(A \to B) \to (\Box A \to \Box B);$$
$$\Box(\Box A \to A) \to \Box A;$$
$$\Box(A \to B) \to (A \rhd B);$$
$$((A \rhd B) \wedge (B \rhd C)) \to (A \rhd C);$$
$$((A \rhd C) \wedge (B \rhd C)) \to ((A \vee B) \rhd C);$$
$$(A \rhd B) \to (\Diamond A \to \Diamond B);$$
$$(\Diamond A) \rhd A;$$
$$(A \rhd B) \to ((A \wedge \Box C) \rhd (B \wedge \Box C)).$$

Thus, *ILM* contains the provability logic *GL* and, therefore, *ILM* $\vdash \Box A \to \Box\Box A$ (see Boolos [2]).

One can show that *ILM* $\vdash \Box A \leftrightarrow (\neg A) \rhd \bot$, which means that $\Box$ can be eliminated from the language of *ILM*.

A finite *Veltman frame* is a system $\langle W, R, \{S_w\}_{w \in W}\rangle$, where W is a finite nonempty set (of "worlds") and R and each S_w are binary relations on W such that the following holds:

1. R is transitive and irreflexive;
2. each S_w is transitive and reflexive;
3. $uS_w v$ only if wRu and wRv;
4. $wRuRv \Longrightarrow uS_w v$;
5. $uS_w vRr \Longrightarrow uRr$.

A finite *Veltman model* is a system

$$\langle W, R, \{S_w\}_{w \in W}, \models\rangle,$$

where $\langle W, R, \{S_w\}_{w\in W}\rangle$ is a finite Veltman frame and $\models$ is a ("forcing") relation between worlds and *ILM*-formulas such that:

- The Boolean connectives are treated in the classical way: $w \not\models \bot$, $w \models A \to B \iff (w \not\models A$ or $w \models B)$, etc.;
- $w \models \Box A \iff$ (for all u, if wRu, then $u \models A$);
- $w \models A \rhd B \iff$ (for all u, if wRu and $u \models A$, then there is v such that $uS_w v$ and $v \models B$).

A formula A is said to be *valid* in a Veltman model $\langle W, R, \{S_w\}_{w\in W}, \models\rangle$, if $w \models A$ for all $w \in W$.

Theorem 2.1 (De Jongh and Veltman [5]) *ILM* $\vdash$ *A iff A is valid in all finite Veltman models.*

3 Arithmetic Preliminaries We fix a theory T containing $I\Sigma_1$. For safety we assume that T is in the language of arithmetic and T is sound, i.e., all its axioms are true (in the standard model of arithmetic). In fact it is easy to adjust our proof of the completeness theorem to the weaker condition of Σ_1-soundness of T.

A *realization* is a function * which assigns an arithmetical sentence p^* to each propositional letter p of the modal language and which is extended to other modal formulas in the following way:

- * commutes with the Boolean connectives: $\bot^* = \bot$, $(A \to B)^* = A^* \to B^*$, etc.;
- $(\Box A)^* = Pr(\ulcorner A^* \urcorner)$;
- $(A \rhd B)^* = Conserv(\ulcorner A^* \urcorner, \ulcorner B^* \urcorner)$,

where $Pr(\ulcorner A^* \urcorner)$ and $Conserv(\ulcorner A^* \urcorner, \ulcorner B^* \urcorner)$ are natural formalizations of "A^* is T-provable" and "T+B^* is Π_1-conservative over T+A^*".

We need to introduce some more notation and terminology.

We will read $\vdash_x F$ as saying that x is the code of some T-proof of the formula F.

We take "Σ_1!" to denote the class of the arithmetical formulas which have an explicit Σ_1 form, i.e., $\exists x F$ for some primitive recursive formula F. And we let "Σ_1" denote the class of the formulas which are T-provably equivalent to some Σ_1!-formula, similarly for Π_1.

Let us fix $\exists y Regwitness(x, y)$ as a natural Σ_1!-formalization of the predicate "x is the code of a true Σ_1!-sentence" such that (T proves that) for each Σ_1!-sentence F, $T \vdash F \leftrightarrow \exists y Regwitness(\ulcorner F \urcorner, y)$.

The existence of the formula $Regwitness(x, y)$ is the only not very trivial—but quite well known (see, e.g., Smorynski [8])—a fact about Σ_1- (Π_1-) sentences that will be used in the arithmetical completeness proof below.

We say that a natural number k is a *regular counterwitness* for a Π_1!-sentence $\forall x F$, if $Regwitness(\ulcorner \exists x \neg F \urcorner, \bar{k})$ is true.

4 The Completeness Theorem

Theorem 4.1 *ILM* $\vdash$ *A iff for any realization* *, $T \vdash A^*$.

The rest of the paper is a proof of this theorem. This proof has a lot of similarity with proofs given in Dzhaparidze [3] and [4], and in Zambella [11]. Just as in [3] and [4], I define here a Solovay function in terms of regular witnesses rather than provability in finite subtheories (as this is done in [1], [7], [11]). Disregarding this difference, my Solovay function is almost the same as the one given in [11]. Both works, unlike [1] or [7], employ finite Veltman models rather than infinite Visser models.

The ($\Longrightarrow$) part of the theorem can be checked by a routine induction on *ILM*-proofs. Here we are going to prove only the ($\Longleftarrow$) part.

Suppose $ILM \nvdash A$. Then, by Theorem 2.1, there is a finite Veltman model $\langle W, R, \{S_w\}_{w\in W}, \models\rangle$ in which A is not valid. We may assume that $W = \{1, \dots, l\}$, 1 is the root of the model in the sense that $1Rw$ for all $1 \neq w \in W$, and $1 \not\models A$.

We define a new frame $\langle W', R', \{S'_w\}_{w\in W'}\rangle$:

$W' = W \cup \{0\}$;

$R' = R \cup \{(0, w) : \ w \in W\}$;

$S'_0 = S_1 \cup \{(1, w) : \ w \in W\}$ and for each $w \in W$, $S'_w = S_w$.

Observe that $\langle W', R', \{S'_w\}_{w\in W'}\rangle$ is a finite Veltman frame.

Following the "traditional" way of arithmetical completeness proofs, we are going to embed this frame into T by means of a Solovay [9] style function $g : \ \omega \to W'$ and sentences Lim_w ($w \in W'$) which assert that w is the limit of g. This function will be defined in such a way that the following basic lemma holds:

Lemma 4.2

a) *T proves that g has a limit in W', i.e., $T \vdash \bigvee\{Lim_r : \ r \in W'\}$.*

b) *If $w \neq u$, then $T \vdash \neg(Lim_w \wedge Lim_u)$.*

c) *If $wR'u$, then $T + Lim_w$ proves that $T \nvdash \neg Lim_u$.*

d) *If $w \neq 0$ and not $wR'u$, then $T + Lim_w$ proves that $T \vdash \neg Lim_u$.*

e) *If $uS'_w v$, then $T + Lim_w$ proves that $T + Lim_v$ is Π_1-conservative over $T + Lim_u$.*

f) *Suppose $wR'u$ and V is a subset of W' such that for no $v \in V$ do we have $uS_w v$. Then $T + Lim_w$ proves that $T + \bigvee\{Lim_v : \ v \in V\}$ is not Π_1-conservative over $T + Lim_u$.*

g) *Lim_0 is true.*

To deduce the main thesis from this lemma, we define a realization * by setting for each propositional letter p,

$$p^* = \bigvee\{Lim_r : \ r \in W, \ r \models p\}.$$

Lemma 4.3 *For any $w \in W$ and any ILM-formula B,*

a) *if $w \models B$, then $T + Lim_w \vdash B^*$;*

b) *if $w \not\models B$, then $T + Lim_w \vdash \neg B^*$.*

Proof: By induction on the complexity of B. If B is atomic, then clause (a) is evident and clause (b) is also clear in view of Lemma 4.2b. The cases when B is a Boolean combination are straightforward; and since $\Box C$ is *ILM*-equivalent to $(\neg C) \rhd \bot$, it is enough to consider only the case when $B = C_1 \rhd C_2$.

Assume $w \in W$. Then we can always write wRx and $xS_w y$ instead of $wR'x$ and $xS'_w y$. Let $\alpha_i = \{r : \ wRr, \ r \models C_i\}$ $(i = 1, 2)$. First we establish that for each $i = 1, 2$,

(*) $\quad T + Lim_w$ proves that $T \vdash C_i^* \leftrightarrow \bigvee\{Lim_r : \ r \in \alpha_i\}$.

We argue in $T + Lim_w$. Since each $r \in \alpha_i$ forces C_i, we have by the induction hypothesis (clause (a)) that for each such r, $T \vdash Lim_r \to C_i^*$, whence $T \vdash \bigvee\{Lim_r : r \in \alpha_i\} \to C_i^*$. Next, according to Lemma 4.2a, $T \vdash \bigvee\{Lim_r : r \in W'\}$ and, according to Lemma 4.2d, T disproves every Lim_r with *not* wRr; consequently, $T \vdash \bigvee\{Lim_r : wRr\}$; at the same time, by the induction hypothesis (clause (b)), C_i^* implies in T the negation of each Lim_r with $r \not\models C_i$. We conclude that $T \vdash C_i^* \to \bigvee\{Lim_r : \ wRr, \ r \models C_i\}$, i.e., $T \vdash C_i^* \to \bigvee\{Lim_r : \ r \in \alpha_i\}$. Thus (*) is proved. Now we continue:

(a) Suppose $w \models C_1 \rhd C_2$. We argue in $T + Lim_w$. By (*), to prove that $T + C_2^*$ is Π_1-conservative over $T + C_1^*$, it is enough to show that $T + \bigvee\{Lim_r : \ r \in \alpha_2\}$ is Π_1-conservative over $T + \bigvee\{Lim_r : \ r \in \alpha_1\}$. Consider an arbitrary $u \in \alpha_1$ (the case with empty α_1 is trivial, for any theory is conservative over $T + \bot$). Since $w \models C_1 \rhd C_2$, there is $v \in \alpha_2$ such that $uS_w v$. Then, by Lemma 4.2e, $T + Lim_v$ is Π_1-conservative over $T + Lim_u$. Then so is $T + \bigvee\{Lim_r : r \in \alpha_2\}$ (which is weaker than $T + Lim_v$). Thus, for each $u \in \alpha_1$, $T + \bigvee\{Lim_r : r \in \alpha_2\}$ is Π_1-conservative over $T + Lim_u$. Clearly this implies that $T + \bigvee\{Lim_r : r \in \alpha_2\}$ is Π_1-conservative over $T + \bigvee\{Lim_r : r \in \alpha_1\}$.

(b) Suppose $w \not\models C_1 \rhd C_2$. Let us then fix an element u of α_1 such that for no $v \in \alpha_2$ do we have $uS_w v$. We argue in $T + Lim_w$. By Lemma 4.2f, $T + \bigvee\{Lim_r : r \in \alpha_2\}$ is not Π_1-conservative over $T + Lim_u$. Then neither is it Π_1-conservative over $T + \bigvee\{Lim_r : r \in \alpha_1\}$ (which is weaker than $T + Lim_u$). This means by (*) that $T + C_2^*$ is not Π_1-conservative over $T + C_1^*$.

Now we can pass to the desired conclusion: since $1 \not\models A$, Lemma 4.3 gives $T \vdash Lim_1 \to \neg A^*$, whence $T \nvdash \neg Lim_1 \Longrightarrow T \nvdash A^*$. But we have $T \nvdash \neg Lim_1$ because, by the Clauses (c) and (g) of Lemma 4.2, this fact is derivable in the sound theory T from the true sentence Lim_0.

Our remaining duty now is to define the function g and prove Lemma 4.2. The Recursion Theorem enables us to define this function simultaneously with the sentences Lim_w (for each $w \in W'$), which, as we have mentioned already, assert that w is the limit of g, and formulas $\Delta_{wu}(y)$ (for each pair (w, u) with $wR'u$), which we define by

$$\Delta_{wu}(y) \equiv \exists t > y(g(t) = \bar{u} \wedge \forall z(y \leq z < t \to g(z) = \bar{w})).$$

Definition 4.4 (of the function g) We define $g(0) = 0$. Assume now $g(y)$ has been defined for every $y \leq x$, and let $g(x) = w$. Then $g(x + 1)$ is defined as follows:

1. Suppose $wR'u$, $n \leq x$ and for all z with $n \leq z \leq x$ we have $g(z) = w$. Then, if $\vdash_x Lim_u \to \neg\Delta_{wu}(\bar{n})$, we define $g(x + 1) = u$.
2. Otherwise suppose $m \leq x$, F is a Π_1!-sentence and the following holds:
   a) F has a regular counterwitness which is $\leq x$;
   b) $\vdash_m Lim_u \to F$;

c) $wS_{g(m)}u$;

d) m is the least number for which such F and u exist, i.e., there are no m' : $m' < m$, world u' and Π_1!-sentence F' satisfying the conditions (a)–(c) when m', u' and F' stand for m, u and F.

Then we define $g(x+1) = u$.

3. In all the remaining cases $g(x+1) = g(x)$.

It is not hard to see that g is primitive recursive. Before we start proving Lemma 4.2, let us agree on some jargon and prove two auxiliary lemmas.

When the transfer from $w = g(x)$ to $u = g(x+1)$ is determined by Definition 4.4.1, we say that at the moment $x+1$ the function g makes (or we make) an *R'-move* from the world w to the world u. If this transfer is determined by Definition 4.4.2, then we say that an *S'-transfer* takes place and call the number m from Definition 4.4.2 the *rank* of this S'-transfer. Sometimes the S'-transfer leads to a new world, but "mostly" it does not, i.e., $(u =)g(x+1) = g(x)(= w)$, and then it is not a move in the proper sense. Those S'-transfers which lead to a new world we call *S'-moves*. As for R'-transfers, they (by irreflexivity of R') always lead to a new world, so we always say "R'-move" instead of "R'-transfer."

In these terms, the formula $\Delta_{wu}(n)$ asserts that beginning from the moment n (but perhaps also before this moment) and until some moment t, we stay at the world w without any motion and then, at the moment t, we move directly to u.

Intuitively, we make an R'-move from w to u, where $wR'u$, in the following situation: since some moment n and up to now we have been staying at the world w, and at the present moment we have reached evidence that $T + Lim_u$ thinks that the first (proper) move which happens after passing the moment n (and thus our next move) cannot lead directly to the world u; then, to spite this belief of $T + Lim_u$, we just move to u.

And the conditions for an S'-transfer from w to u can be described as follows: We are staying at the world w and by the present moment we have reached evidence that $T + Lim_u$ proves a false Π_1!-sentence F. This evidence consists of two components: (1) a regular counterwitness, which indicates that F is false, and (2) the rank m of the transfer, which indicates that $T + Lim_u \vdash F$. Then, as soon as $wS_{g(m)}u$, the next moment we must be at u (move to u, if $u \neq w$, and remain at w, if $u = w$); if there are several possibilities of this transfer, we choose the one with the least rank. Besides, the necessary condition for an S'-transfer is that in the given situation an R'-move is impossible.

Lemma 4.5 $(T \vdash:)$ *For each natural number m and each $w \in W'$, $T + Lim_w$ proves that no S'-transfer to w can have rank which is less than m.*

Proof: Note that "the rank of an S'-transfer is $< m$" means that $T + Lim_w$ proves a false Π_1!-sentence F (i.e., one with a regular counterwitness) and the code of this proof (i.e., of the T-proof of $Lim_w \to F$) is smaller than m. But the number of all Π_1!-sentences with such short proofs is finite, and as $T + Lim_w$ proves each of them, it also proves that none of these sentences has a regular counterwitness (recall our assumptions about the formula $Regwitness(x, y)$).

Lemma 4.6 ($T \vdash$:) *If $g(x)R'w$, then for all $y \le x$, $g(y)R'w$.*

Proof: Suppose $g(x)R'w$ and $y \le x$. We proceed by induction on $n = x - y$. If $y = x$, we are done. Suppose now $g(y+1)R'w$. If $g(y) = g(y+1)$, we are done. If not, then at the moment $y + 1$ the function makes either an R'-move or an S'-move. In the first case we have $g(y)R'g(y+1)$ and, by transitivity of R', $g(y)R'w$; in the second case we have $g(y)S'_v g(y+1)$ for some v, and the desired thesis then follows from the Property 5 of Veltman frames.

Proof: (of Lemma 4.2) In each case below, except (g), we reason in T.

(a) First observe that there is z such that for all $z' \ge z$, not $g(z')R'g(z'+1)$.

Indeed, suppose this is not the case. Then, by Lemma 4.6, for all z there is z' with $g(z)R'g(z')$. This means that there is an infinite (or "sufficiently long") chain $w_1 R' w_2 R' \ldots$, which is impossible because W' is finite and R' is transitive and irreflexive.

So let us fix this number z. Then we never make an R'-move after the moment z. We claim that S'-moves can also take place at most a finite number of times (whence it follows that g has a limit and this limit is, of course, one of the elements of W').

Indeed, let x be an arbitrary moment after z at which we make an S'-move, and let m be the rank of this move. Taking into account reflexivity of the relations S_w, a little analysis of the Condition 4.4.2 convinces us that the rank of each next S'-move is less than that of the previous one, so S'-moves can take place at most m times after passing x.

(b) Clearly g cannot have two different limits w and u.

(c) Assume w is the limit of g and $wR'u$. Let n be such that for all $x \ge n$, $g(x) = w$. We need to show that $T \nvdash \neg Lim_u$. Suppose this was the case. Then $T \vdash Lim_u \to \neg\Delta_{wu}(\bar{n})$ and, since every provable formula has arbitrary long proofs, there is $x \ge n$ such that $\vdash_x Lim_u \to \neg\Delta_{wu}(\bar{n})$. But then, according to Definition 4.4.1, we must have $g(x+1) = u$, which, as $u \ne w$ (by irreflexivity of R'), is a contradiction.

(d) Assume $w \ne 0$, w is the limit of g and not $wR'u$.

If $u = w$, then (since $w \ne 0$) there is x such that $g(x) = v \ne u$ and $g(x+1) = u$. This means that at the moment $x + 1$ we make either an R'-move or an S'-move. In the first case we have $T \vdash Lim_u \to \neg\Delta_{vu}(\bar{n})$ for some n for which, as it is easy to see, the Σ_1!-sentence $\Delta_{vu}(\bar{n})$ is true, whence, by Σ_1!-completeness, $T \vdash \neg Lim_u$. And if an S'-move is the case, then again $T \vdash \neg Lim_u$ because $T + Lim_u$ proves a false (with a $\le x$ regular counterwitness) Π_1!-sentence.

Suppose now $u \ne w$. Let us fix a number z with $g(z) = w$. Since g is primitive recursive, T proves that $g(z) = w$.

Now we argue in $T + Lim_u$: Since u is the limit of g and $g(z) = w \ne u$, there is a number x with $x \ge z$ such that $g(x) \ne u$ and $g(x+1) = u$. Since not $(w =)g(z)R'u$, we have by Lemma 4.6 that

(*) For each y with $z \le y \le x$, not $g(y)R'u$.

In particular, not $g(x)R'u$ and the transfer from $g(x)$ to $g(x+1)(= u)$ can be determined only by Definition 4.4.2. Then (*) together with the Property 3 of Veltman frames and Definition 4.4.2c, implies that the rank of this S'-move is less than z, which, by Lemma 4.5, is a contradiction.

Thus, $T + Lim_u$ is inconsistent, i.e., $T \vdash \neg Lim_u$.

(e) Assume $uS'_w v \neq u$ (the case $v = u$ is trivial). Suppose w is the limit of g, F is a Π_1-sentence and $T \vdash_z Lim_v \to F$. We may suppose that $F \in \Pi_1!$ and that z is sufficiently large, namely, $g(z) = w$. Fix this z. We need to show that $T + Lim_u \vdash F$.

We argue in $T + Lim_u$. Suppose not F. Then there is a regular counterwitness c for F. Let us fix a number $x > z, c$ such that $g(x) = g(x+1) = u$ (as u is the limit of g, such a number exists). Then, according to 4.4.2, the only reason for $g(x+1) = u \neq v$ can be that we make an S'-transfer from u to u and the rank of this transfer is less than z, which, by Lemma 4.5, is not the case. We therefore conclude that F (is true).

(f) Assume w is the limit of g, $wR'u$, $V \subseteq W'$ and for each $v \in V$, not , $uS'_w v$.

Let n be such that for all $z \geq n$, $g(z) = w$. By the primitive recursiveness of g, T proves that $g(n) = w$. By 4.4.1, $T + Lim_u \nvdash \neg\Delta_{wu}(\bar{n})$. So, as $\neg\Delta_{wu}(\bar{n})$ is a Π_1-sentence, in order to prove that $T + \bigvee\{Lim_v : \ v \in V\}$ is not Π_1-conservative over $T + Lim_u$, it is enough to show that for each $v \in V$, $T + Lim_v \vdash \neg\Delta_{wu}(\bar{n})$. Let us fix any $v \in V$. According to our assumption, not $uS'_w v$ and, by reflexivity of S'_w, $u \neq v$.

We now argue in $T + Lim_v$. Suppose, for a contradiction, that $\Delta_{wu}(n)$ holds, i.e., there is $t > n$ such that $g(t) = u$ and for all z with $n \leq z < t$, $g(z) = w$. As v is the limit of g and $v \neq u$, there is $t' > t$ such that $g(t'-1) \neq v$ and at the moment t' we arrive to v to stay there for ever. Let then $x_0 < \ldots < x_k$ be all the moments in the interval $[t, t']$ at which R'- or S'-moves take place, and let $u_0 = g(x_0), \ldots, u_k = g(x_k)$. Thus $t = x_0$, $t' = x_k$, $u = u_0$, $v = u_k$ and $u_0, \ldots, u_k$ is the route of g after departing from w (at the moment t).

Now let j be the least number among $1, \ldots, k$ such that for all $j \leq i \leq k$, not $u_0 R' u_i$. Note that such a j does exist because at least $j = k$ satisfies this condition (otherwise, if $(u =)u_0 R' u_k (= v)$, Property 4 of Veltman frames would imply $uS'_w v$).

Note also that for each i with $j \leq i \leq k$, the move to u_i cannot be an R'-move. Indeed, otherwise we must have $u_{i-1} R' u_i$, whence, by Lemma 4.6, $u_0 R' u_i$, which is impossible for $i \geq j$.

Thus, beginning from the moment x_j (inclusive), each move is an S'-move. Moreover, for each i with $j \leq i \leq k$, the rank of the S'-move to u_i is less than x_0. For otherwise Property 3 of Veltman frames together with Lemma 4.6 and Definition 4.4.2c would entail that $u_0 R' u_i$. On the other hand, since consecutive S'-moves decrease the rank (as we noted in the proof of (a) above) and since the rank of the S'-move to u_k cannot be less than n (Lemma 4.5), we conclude that for each i with $j \leq i \leq k$, the rank of the S'-move to u_i is in the interval $[n, x_0 - 1]$. But the value of g in this interval is w, and by Definition 4.4.2c this means that $u_{j-1} S'_w u_j S'_w \ldots S'_w u_k$. At the same time, we have either $u_0 = u_{j-1}$ or $u_0 R' u_{j-1}$. In both cases we then have $u_0 S'_w u_{j-1}$ (in the first case by reflexivity of S'_w and in the second case by the Property 4 of Veltman frames), whence, by transitivity of S'_w, $u_0 S'_w u_k$, i.e., $uS'_w v$, which is a contradiction.

Thus we can conclude that $T + Lim_v \vdash \neg\Delta_{wu}(\bar{n})$.

(g) By Lemma 4.2a, as T is sound, one of the Lim_w $(w \in W')$ is true. Since for no w do we have $wR'w$, Lemma 4.2d means that each Lim_w, except Lim_0, implies in T its own T-disprovability and therefore is false. Consequently, Lim_0 is true. This completes the proof of Lemma 4.2.

This in turn completes the proof of Theorem 4.1.

REFERENCES

[1] Berarducci, A., "The interpretability logic of Peano Arithmetic," *The Journal of Symbolic Logic*, vol. 55 (1990), pp. 1059–1089. Zbl 0725.03037 MR 92f:03066 1, 4, 4

[2] Boolos, G., *The Logic of Provability*, Cambridge University Press, Cambridge, 1993. Zbl 0891.03004 MR 95c:03038 2

[3] Dzhaparidze (Japaridze), G., "The logic of linear tolerance," *Studia Logica*, vol. 51 (1992), pp. 249–277. Zbl 0769.03010 MR 94e:03061 4, 4

[4] Dzhaparidze (Japaridze), G., "A generalized notion of weak interpretability and the corresponding modal logic," *Annals of Pure and Applied Logic*, vol. 61 (1993), pp. 113–160. Zbl 0791.03032 MR 94d:03030 4, 4

[5] de Jongh, D., and F. Veltman, "Provability logics for relative interpretability," pp. 31–42 in *Mathematical Logic*, edited by P. Petkov, Plenum Press, New York, 1990. Zbl 0794.03026 MR 92d:03011 1, 2.1

[6] Hájek, P., and F. Montagna, "The logic of Π_1-conservativity," *Archive for Mathematical Logic*, vol. 30 (1990), pp. 113–123. Zbl 0713.03007 MR 92a:03024 1

[7] Shavrukov, V., *The logic of relative interpretability over Peano Arithmetic* (in Russian). Preprint No. 5, Steklov Mathematical Institute, Moscow, 1988. 1, 4, 4

[8] Smorynski, C., "The incompleteness theorems," pp. 821–865 in *Handbook of Mathematical Logic*, edited by J. Barwise, North-Holland, Amsterdam, 1977. Zbl 0443.03001 MR 56:15351 3

[9] Solovay, R., "Provability interpretations of modal logic," *Israel Journal of Mathematics*, vol. 25 (1976), pp. 287–304. Zbl 0352.02019 MR 56:15369 4

[10] Visser, A., "Interpretability logic," pp. 175–209 in *Mathematical Logic*, edited by P. Petkov, Plenum Press, New York, 1990. Zbl 0793.03064 MR 93k:03022 1

[11] Zambella, D., "On the proofs of arithmetical completeness of interpretability logic," *Notre Dame Journal of Formal Logic*, vol. 33 (1992), pp. 542–551. Zbl 0788.03027 MR 93k:03022 4, 4, 4

Department of Computer and Information Science
University of Pennsylvania
200 South 33rd Street
Philadelphia, PA 19104–6389
email: giorgi@gradient.cis.upenn.edu