A short note on essentially Σ_1 sentences

Franco Montagna[(1)], Duccio Pianigiani[(2)]
[(1)]Department of Mathematics and Computer Science
University of Siena, Italy
[(2)]Department of Philosophy
University of Siena, Italy

Abstract David Guaspari [G] conjectured that a modal formula is *essentially* Σ_1 (i.e., is is Σ_1 under any arithmetical interpretation), if and only if it is provably equivalent to a disjunction of formulas of the form $\Box B$. This conjecture was proved first by A. Visser. Then, in [DJP], the authors characterized essentially Σ_1 formulas of languages with witness comparisons using the interpretability logic **ILM**. In this note we give a similar characterization for formulas with a binary operator interpreted as interpretability in a finitely axiomatizable extension of $\mathbf{I\Delta_0} + \mathbf{Supexp}$ and we address a similar problem for $\mathbf{I\Delta_0} + \mathbf{Exp}$.

Keywords: First-order fragments of arithmetic, Interpretability logic.

1 Introduction and statement of the problem

The problem investigated in this note involved the present authors and Dick De Jongh in the 80's. In these years, De Jongh, Veltman, Visser and other people invented Interpretability Logic, [V1], [V2], [DV], and the first author contributed partially to the axiomatization of the interpretability logic of Peano Arithmetic. Moreover, in [DJP] Dick De Jongh and the second author found a characterization of formulas of Provability Logic **L** (cf [Sm]) and of Guaspari and Solovay's system **R** (cf [GS]) which are essentially Σ_1 with respect to any Σ_1-sound r.e. extension of $\mathbf{I\Sigma_1}$, thus solving an open problem by Guaspari [G]. At that time, we had an enjoyable cooperation with the Dutch research group in Provability Logic, and with Dick De Jongh in particular. His visits to Siena and our visits to Amsterdam and to Utrecht were extremely useful for us, and we learned a lot of things during them. Then, for some reasons which we are not going to explain here, we left Provability Logic. Thus, when Veltman invited us to contribute to the book for Dick, we had two kinds of reactions: on one hand, we felt happy to return in our mind to the good old times, and on the other hand we felt worried because we realized that we forgot too many things for being able to find a really good result. Thus, our contribution is very modest: we could say that De Jongh, Hájek, Veltman, Visser and other people gave a great work to grow apple trees, and we only contributed to eat apples. Still, we did our best, therefore we hope that Dick will appreciate our efforts.

Definition 1.1 (See [DJP]) Let $\mathcal{L}$ be a logic with a collection of arithmetical interpretations into an r.e. extension $\mathbf{T}$ of $\mathbf{I\Delta_0} + \mathbf{Exp}$. We say that a formula A is *essentially* Σ_1 with respect to $\mathbf{T}$, if $A^\star$ is provably equivalent to a Σ_1 formula for any arithmetical interpretation $\star$ of $\mathcal{L}$ in $\mathbf{T}$.

In [G], David Guaspari formulated the following conjectures:

- *A formula A of Provability Logic $\mathbf{L}$ (cf [Sm]) is essentially Σ_1 with respect to Peano Arithmetic $\mathbf{PA}$ iff A is provably equivalent in $\mathbf{L}$ to a disjunction of formulas of the form $\Box B$*

- *A formula A of the system $\mathbf{R}$ of Guaspari and Solovay (cf [GS]), is essentially Σ_1 with respect to $\mathbf{PA}$ iff it is provably equivalent in $\mathbf{R}$ to a lattice combination of formulas of either form $\Box B$, $\Box B \preceq \Box C$, or $\Box B \prec \Box C$*

The first question received a positive answer by Albert Visser (cf [V3]). He also suggested to consider the interpretability logic systems **ILM** and **IRM** (cf [DV] and [HM]) to face this problem. Then in [DPJ] De Jongh and Pianigiani extended the result to the system $\mathbf{R}$, with reference to any r.e. Σ_1-sound extension of $\mathbf{I\Sigma_1}$. More precisely, they showed the following:

Theorem 1.2 *Let $\mathbf{T}$ be any Σ_1 sound r.e. extension of $\mathbf{I\Sigma_1}$, and let A be any formula of $\mathbf{L}$. The following are equivalent:*

(a) A is essentially Σ_1 with respect to $\mathbf{T}$.

(b) $\mathbf{ILM} \vdash (p \rhd q) \to ((p \wedge A) \rhd (q \wedge A))$ for all p, q not occurring in A.

(c) A is provably equivalent to a disjunction of formulas of the form $\Box B$.

Moreover, let A be a formula of $\mathbf{R}$. The following are equivalent:

(d) A is essentially Σ_1 with respect to $\mathbf{T}$.

(e) $\mathbf{IRM} \vdash (p \rhd q) \to ((p \wedge A) \rhd (q \wedge A))$ for all p, q not occurring in A.

(f) A is a lattice combination of formulas of either form $\Box B$, $\Box B \preceq \Box C$ or $\Box B \prec \Box C$

A characterization of **ILM** formulas which are essentially Σ_1 with respect to any essentially reflexive extension of $\mathbf{I}\Sigma_1$ was recently given in [GJ].

In this note we show that a characterization of this kind can be obtained also for formulas of the logic **ILP** (cf [V1]), with respect to any finitely axiomatizable Σ_1-sound extension of $\mathbf{I\Delta_0} + \mathbf{Supexp}$. More precisely, we prove:

Theorem 1.3 *Let A be a formula of the language of* **ILP***, and let* **T** *be a finitely axiomatizable Σ_1-sound extension of* $\mathbf{I\Delta_0} + \mathbf{Supexp}$*. Then A is essentially Σ_1 with respect to* **T** *iff it is provably equivalent in* **ILP** *to a lattice combination of formulas of the form $C \rhd D$.*

In the last section, we prove that the same characterization does not extend to $\mathbf{I\Delta_0} + \mathbf{Exp}$, and we formulate a conjecture about essentially Σ_1 **ILP**-formulas with respect to $\mathbf{I\Delta_0} + \mathbf{Exp}$.

2 Preliminary notions

2.1 Theories of arithmetic

In the sequel, $\mathbf{I\Delta_0}$ denotes Robinson's system **Q** plus induction restricted to bounded arithmetical formulas. $\mathbf{I\Delta_0} + \mathbf{Exp}$ is $\mathbf{I\Delta_0}$ plus the totality of exponentiation, and $\mathbf{I\Delta_0} + \mathbf{Supexp}$ is $\mathbf{I\Delta_0}$ plus the totality of superexponentiation. For $n > 0$ $\mathbf{I\Sigma_n}$ denotes **Q** plus induction restricted to Σ_n formulas. In the sequel, given an arithmetical theory **T**, $\Box_{\mathbf{T}}$ denotes a standard proof predicate for **T**, and $\Box^{cf}_{\mathbf{T}}$ denotes a standard predicate expressing cut-free provability in **T**. We use the abbreviations $\Box_{\mathbf{Exp}}$ and $\Box_{\mathbf{Supexp}}$ ($\Box^{cf}_{\mathbf{Exp}}$ and $\Box^{cf}_{\mathbf{Supexp}}$ respectively) to denote standard (cut-free respectively) proof predicates for $\mathbf{I\Delta_0} + \mathbf{Exp}$ and for $\mathbf{I\Delta_0} + \mathbf{Supexp}$. Moreover $\Diamond_{\mathbf{T}}$ and $\Diamond^{cf}_{\mathbf{T}}$ will denote $\neg\Box_{\mathbf{T}}\neg$ and $\neg\Box^{cf}_{\mathbf{T}}\neg$ respectively.

2.2 Interpretability logics

Interpretability logics are modal logics expressing (provability and) interpretability, see [V2] for a general survey. The languages of such logics are built-up from propositional variables, $\bot$, connectives, and the binary operator $\rhd$. The intended meaning of $A \rhd B$ is: $\mathbf{T} + B$ is interpretable in $\mathbf{T} + A$. We write $\Box A$ for $(\neg A) \rhd \bot$, and $\Diamond A$ for $\neg\Box\neg A$.

Definition 2.1 The basic interpretability logic **IL** (cf [V1]) has the following axioms and rules:

1. The axioms and rules of Provability Logic **L**.
2. $(A \rhd B) \wedge (B \rhd C) \to (A \rhd C)$
3. $(A \rhd B) \wedge (C \rhd B) \to ((A \vee C) \rhd B)$
4. $(A \rhd B) \to (\Diamond A \to \Diamond B)$
5. $(\Diamond A) \rhd A$

Let **T** be an r.e. theory of arithmetic. An *interpretation of* **IL** *in* **T** is a map * from **IL**-formulas into **T**-sentences which commutes with all connectives and such that $\bot^\star = \bot$, and for every **IL**-formulas A and B, $(A \rhd B)^\star$ is the formalization of: $\mathbf{T} + B^\star$ *is intepretable in* $\mathbf{T} + A^\star$. A modal formula A is *valid in* **T** iff for every interpretation $\star$ in **T**, $\mathbf{T} \vdash A^\star$.

Every theorem of **IL** is valid in every Σ_1-sound extension of $\mathbf{I\Delta_0} + \mathbf{Exp}$. There are two modal principles which are valid in some theories and not valid in other theories.

- (M) is the principle $(A \rhd B) \to ((A \wedge \Box C) \rhd (B \wedge \Box C))$.
- (P) is the principle $(A \rhd B) \to \Box(A \rhd B)$.

ILM is **IL** plus (M), and **ILP** is **IL** plus (P). It is well-known ([V1] and [DV]) that (P) is valid in every finitely axiomatizable extension of $\mathbf{I\Delta_0} + \mathbf{Exp}$, and (M) is valid in every essentially reflexive extension of $\mathbf{I\Delta_0} + \mathbf{Exp}$.

2.3 Kripke semantics for ILP

A complete Kripke-style semantics for **ILP** is the following one.

Definition 2.2 (cf [V1]). A (Carlson-style) *finite Friedman structure* is a system $\langle K, b, D, Q\rangle$ such that:

(i) K is a finite set, whose elements are called *nodes*, $D \subseteq K$, and $b \in D$ (hence, $b \in K$, and $D, K \neq \emptyset$. b is called *the root* of the structure).

(ii) Q is a binary treelike relation on K (i.e., Q is irreflexive, transitive, and xQz and yQz implies $x = y$).

(iii) For all $x \in K \setminus \{b\}$, bQx.

(iv) If $x, y \in D$ and xQy, then there is $z \in K$ such that $xQzQy$.

A *forcing* on $\langle K, b, D, Q\rangle$ is a binary relation $\Vdash$ between K and the set of **ILP**-formulas such that, writing $x \Vdash A$ for $(x, A) \in \Vdash$, the following conditions hold:

(v) The usual conditions on forcing in Kripke models for Boolean connectives and for $\bot$.

(vi) $x \Vdash A \rhd B$ iff for all $u, v \in K$, if $xQuQv$, $v \in D$ and $v \Vdash A$, then there is $w \in D$ such that uQw and $w \Vdash B$.

A (Carlson-style) *finite Friedman model* is a system $\langle K, b, D, Q, \Vdash\rangle$ such that $\langle K, b, D, Q\rangle$ is a finite Friedman structure and $\Vdash$ is a forcing on it.

Notice that according to this definition, $x \Vdash \Box A$ iff for all u, z if $xQuQz$ and $z \in D$, then $z \Vdash A$. Thus letting xRz iff $z \in D$ and there is $u \in K$ with $xQuQz$, we obtain the usual forcing condition for $\Box$: $x \Vdash \Box A$ iff for all z, if xRz then $z \Vdash A$.

2.4 Kripke completeness and arithmetical completeness of ILP

In [V1], Albert Visser proves the following:

Theorem 2.3 *Let A be a* **ILP***-formula, and let* **T** *be a finitely axiomatizable Σ_1-sound extension of* $\mathbf{I\Delta_0} + \mathbf{Exp}$*. The following are equivalent:*

(a) $\mathbf{ILP} \vdash A$.

(b) For every finite Friedman model $\langle K, b, D, Q, \Vdash \rangle$, $b \Vdash A$.

(c) For every interpretation $\star$ in **T***,* $\mathbf{T} \vdash A^\star$.

Remark 1 In order to prove (c)⇒(b), Visser starts from a finite Friedman model $\langle K, b, D, Q, \Vdash \rangle$ such that $b \Vdash \neg A$ and constructs, for every $i \in K$, a sentence L_i and an interpretation * in **T** such that:

(i) For all $i \in K$, $\mathbf{I\Delta_0} + \mathbf{Exp} + L_i$ is consistent, and for all $i \in D$, $\mathbf{T} + L_i$ is consistent.

(ii) For every subformula C of A and for all $i \in K$, if $i \Vdash C$, then $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash L_i \to C^\star$, and if $i \vdash \neg C$, then $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash L_i \to \neg C^\star$.

(iii) For all $i, j \in K$, if iQj, then $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash L_i \to \neg \Box^{cf}_{\mathbf{Exp}}(\neg L_j)$.

3 Essentially Σ_1 formulas in finitely axiomatizable extensions of $\mathbf{I\Delta_0} + \mathbf{Supexp}$

Definition 3.1 In the sequel $C\Sigma$ denotes the set of **ILP**-formulas which are finite conjunctions of formulas of the form $B \rhd C$, and Σ denotes the set of **ILP**-formulas which are provably equivalent to disjunctions of formulas in $C\Sigma$. Moreover, $N\Sigma$ denotes the set of finite conjunctions of literals (propositional atoms and negations of propositional atoms) and of formulas of the form $\neg(B \rhd C)$.

Definition 3.2 We say that a **ILP**-formula A is in *disjunctive normal form* (*DNF* for short) if $A = \bigvee_{i=1}^n A_i$, where each A_i is either in $C\Sigma$ or of the form $S_i \wedge B_i$, where $S_i \in C\Sigma$ and $B_i \in N\Sigma$.

Clearly, every **ILP**-formula is provably equivalent to a formula in DNF.

Definition 3.3 Say that a **ILP**-formula $A = \bigvee_{i=1}^n A_i$ in DNF is *reducible* if there is $i \leq n$ such that:

- $A_i = S_i \wedge B_i$, with $S_i \in C\Sigma$ and $B_i \in N\Sigma$.

- Letting $R_i = \bigvee_{j \neq i} A_j$, one has either $\mathbf{ILP} \vdash A_i \to R_i$, or $\mathbf{ILP} \vdash (S_i \wedge \neg B_i) \to R_i$.

We say that a formula A in DNF is *irreducible* if it is not reducible.

As noted in [DJP], if $\mathbf{ILP} \vdash A_i \to R_i$ then A is provably equivalent to R_i, and if $\mathbf{ILP} \vdash (S_i \wedge \neg B_i) \to R_i$, then A is provably equivalent to $S_i \wedge R_i$. In other words, if A is reducible, then A is provably equivalent to a shorter formula. It follows:

Lemma 3.4 *Any* **ILP** *formula is provably equivalent to an irreducible formula. Moreover, if $A \notin \Sigma$, then A can be written in DNF as $A = \bigvee_{i=1}^{n} A_i$, where $\bigvee_{i=1}^{n} A_i$ is irreducible, $A_1 = S_1 \wedge B_1$, $S_1 \in C\Sigma$, $B_1 \in N\Sigma$, and both $\mathbf{ILP} + S_1 + B_1 + \neg R_1$ and $\mathbf{ILP} + S_1 + \neg B_1 + \neg R_1$ are consistent.*

■

Lemma 3.5 *Suppose $A \notin \Sigma$. Then there is a finite Friedman model $\langle K, b, D, Q, \Vdash \rangle$ such that:*

(i) $b \Vdash A$.

(ii) There is a node $x \in K$ such that $x \Vdash \neg A$.

Proof. By Lemma 3.4, we can write A in DNF as $A = \bigvee_{i=1}^{n} A_i$ where $A_1 = S_1 \wedge B_1$, $S_1 \in C\Sigma$, $B_1 \in N\Sigma$, and both $\mathbf{ILP} + S_1 + B_1 + \neg R_1$ and $\mathbf{ILP} + S_1 + \neg B_1 + \neg R_1$ are consistent. By Theorem 2.3, there are finite Friedman models $\langle K_1, b_1, D_1, Q_1, \Vdash_1 \rangle$, $\langle K_2, b_2, D_2, Q_2, \Vdash_2 \rangle$ such that $b_1 \Vdash_1 S_1 \wedge B_1 \wedge \neg R_1$, and $b_2 \Vdash_2 S_1 \wedge \neg B_1 \wedge \neg R_1$. We can suppose without loss of generality that $K_1 \cap K_2 = \emptyset$. Now let $b \notin K_1 \cup K_2$, and construct a model $\langle K, b, D, Q, \Vdash \rangle$ with root b as follows:

- $K = K_1 \cup K_2 \cup \{b\}$.
- $Q = Q_1 \cup Q_2 \cup \{(b, x) : x \in K_1 \cup K_2\}$.
- $D = (D_1 \cup D_2 \cup \{b\}) \setminus \{b_1, b_2\}$.
- For every propositional variable p and for every $x \in K$, $x \Vdash p$ iff either $x \in K_1$ and $x \Vdash_1 p$, or $x \in K_2$ and $x \Vdash_2 p$, or $x = b$ and $b_1 \Vdash_1 p$.

It is easily verified that $\langle K, b, D, Q, \Vdash \rangle$ is really a finte Friedman model. Clearly, for $x \in K_i$ $(i = 1, 2)$ and for every formula C, $x \Vdash C$ iff $x \Vdash_i C$. Thus $b_2 \Vdash S_1 \wedge \neg B_1 \wedge \neg R_1$, therefore $b_2 \Vdash \neg A$. So, it remains to prove that $b \Vdash A$. Clearly, it is sufficient to prove that $b \Vdash S_1 \wedge B_1$.

We first prove that $b \Vdash S_1$. Consider any conjunct in S_1 of the form $E \rhd F$. Then $b_1 \Vdash_1 E \rhd F$, and $b_2 \Vdash_2 E \rhd F$. Suppose that $bQuQz$, $z \in D$ and $z \Vdash E$. Suppose e.g. $u \in K_1$. Then $z \in D_1$, $uQ_1 z$ and $z \Vdash_1 E$. Distinguish two cases:

(i) If $u \neq b_1$, then $b_1 Q_1 u Q_1 z$, $z \in D_1$ and $z \Vdash_1 E$. Since $b_1 \Vdash_1 E \rhd F$, there is $v \in D_1$ such that $u Q_1 v$ and $v \Vdash_1 F$. Hence $bQuQv$, $v \in D$ and $v \Vdash F$. Summing-up, if $bQuQz$, $z \in D$ and $z \Vdash E$, there is $v \in D$ such that uQv and $v \Vdash F$. In other words, $b \Vdash E \rhd F$.

(ii) If $u = b_1$, then since $b_1 \in D_1$, by condition (iv) in Definition 2.2, there is $w \in K_1$ with $b_1 Q_1 w Q_1 z$. Since $b_1 \Vdash_1 E \rhd F$, there is $v \in D_1$ such that wQv and $v \Vdash_1 F$. Hence $bQu = b_1 Qv$, $v \in D$, and $v \Vdash F$. Once again, $b \Vdash E \rhd F$.

Now we prove that $b \Vdash B_1$. Let H be a conjunct in B_1. If H is a Boolean formula, then $b \Vdash H$, because $b_1 \Vdash_1 H$, and b and b_1 force the same Boolean formulas. If $H = \neg(E \rhd F)$, then it cannot be the case that $b \Vdash E \rhd F$, becasue Σ formulas are upwards preserved in every Friedman model, and $b_1 \Vdash \neg(E \rhd F)$.

Summing-up, $b \Vdash S_1 \wedge B_1$, $b \Vdash A$, and the claim is proved. ■

Theorem 3.6 *Let* $\mathbf{T}$ *be any finitely axiomatizable and* Σ_1*-sound extension of* $\mathbf{I\Delta_0} + \mathbf{Supexp}$*, and let* A *be a* **ILP***-formula. Then* A *is essentially* Σ_1 *with respect to* $\mathbf{T}$ *iff* $A \in \Sigma$*, i.e., iff* A *is provably equivalent in* **ILP** *to a lattice combination of formulas of the form* $B \rhd C$*.*

Proof. The right-to-left implication is trivial, as interpretability between finitely axiomatizable theories is an r.e. relation. Conversely, suppose $A \notin \Sigma$. Then by Lemma 3.5, there are a finite Friedman model $\langle K, b, D, Q, \Vdash \rangle$ with root b and a node x such that $b \Vdash A$, and $x \Vdash \neg A$. Now by Theorem 2.3 and by Remark 1, we have:

(a) $\mathbf{T} + L_b$ is consistent.

(b) $\mathbf{I\Delta_0} + \mathbf{Exp} + L_b \vdash A^\star$.

(c) $\mathbf{I\Delta_0} + \mathbf{Exp} + L_b \vdash \neg\Box^{cf}_{\mathbf{Exp}}(\neg L_x)$.

(d) $\mathbf{I\Delta_0} + \mathbf{Exp} + L_x \vdash \neg A^\star$.

We claim that $A^\star$ is not provably equivalent to a Σ_1-formula. Suppose it is. Then by Σ_1-completeness, $\mathbf{I\Delta_0} + \mathbf{Exp} + L_b \vdash A^\star \rightarrow \Box^{cf}_{\mathbf{Exp}}(A^\star)$. This is impossible, because $\mathbf{I\Delta_0} + \mathbf{Exp} + L_b$ is consistent by (a), and (b), (c) and (d) together give

$$\mathbf{I\Delta_0} + \mathbf{Exp} + L_b \vdash A^\star \wedge \neg\Box^{cf}_{\mathbf{Exp}}(A^\star).$$

■

4 Formulas of ILP which are essentially Σ_1 with respect to $\mathbf{I\Delta_0} + \mathbf{Exp}$

In the sequel, if A, B are formulas of $\mathbf{I\Delta_0} + \mathbf{Exp}$, then $A \rhd_{\mathbf{Exp}} B$ denotes the formalization of: $\mathbf{I\Delta_0} + \mathbf{Exp} + B$ *is interpretable into* $\mathbf{I\Delta_0} + \mathbf{Exp} + A$. In [V1], Visser proves:

Theorem 4.1 $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash (A \rhd_{\mathbf{Exp}} B) \leftrightarrow \Box^{cf}_{\mathbf{Exp}}(\Diamond^{cf}_{\mathbf{Exp}}(A) \rightarrow \Diamond^{cf}_{\mathbf{Exp}}(B))$.

We define a (not faithful) interpretation $^\circ$ of **ILP** into **L** (where the necessity and the possibility operators are denoted by Δ and by ∇ respectively) in the following inductve way:

$p^\circ = p$ for every atomic formula p.

$^\circ$ commutes with Boolean connectives.

$(A \rhd B)^\circ = \Delta(\nabla(A^\circ) \rightarrow \nabla(B^\circ))$.

Since the provability logic of $\Box^{cf}_{\mathbf{Exp}}$ is just provability Logic **L**, we have:

Theorem 4.2 *Let A be any* **ILP***-formula. The following are equivalent:*

(i) $\mathbf{L} \vdash A^\circ$.

(ii) For every interpretation $\star$ in $\mathbf{I\Delta_0} + \mathbf{Exp}$, $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash A^\star$.

Note that the operator $\Box$ representing $\Box_{\mathbf{Exp}}$ is defined by means of the interpretability operator as usual, i.e. $\Box(A) = (\neg A) \rhd \bot$. An easy computation shows that $\Box(A)^\circ$ is provably equivalent to $\Delta\Delta(A^\circ)$. Hence, so to speak, Δ is the square root of $\Box$.

M.Kalsbeck [K] proved that **ILP** is not complete with respect to $\mathbf{I\Delta_0} + \mathbf{Exp}$, that is, there are formulas A such that $\mathbf{ILP} \nvdash A$ and $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash A^\star$ for every interpretation $\star$ in $\mathbf{I\Delta_0} + \mathbf{Exp}$. As a consequence, we obtain:

Corollary 4.3 *The set of* **ILP** *formulas which are essentially Σ_1 with respect to* $\mathbf{I\Delta_0} + \mathbf{Exp}$ *is a proper superset of Σ. Thus there are* **ILP***-formulas which are essentially Σ_1 with respect to* $\mathbf{I\Delta_0} + \mathbf{Exp}$*, but not with respect to* $\mathbf{I\Delta_0} + \mathbf{Supexp}$.

Proof. Let A such that $\mathbf{ILP} \nvdash A$ and $\mathbf{I\Delta_0} + \mathbf{Exp} \vdash A^\star$ for every interpretation $\star$ in $\mathbf{I\Delta_0} + \mathbf{Exp}$. Then clearly $\mathbf{ILP} \nvdash \neg A$, therefore there are finite Friedman models $\langle K_1, b_1, D_1, Q_1, \Vdash_1\rangle$, $\langle K_2, b_2, D_2, Q_2, \Vdash_2\rangle$ such that $b_1 \Vdash_1 A$, and $b_2 \Vdash_2 \neg A$. Now we can construct a finite Friedman model $\langle K, b, D, Q, \Vdash\rangle$ with a new root b along the lines of the proof of Lemma 3.5. Then $b_1 \Vdash A$, $b_2 \Vdash \neg A$, therefore, independently of whether $b \Vdash A$ or $b \Vdash \neg A$, we have that either A or $\neg A$ is not upwards preserved, therefore either A or $\neg A$ is not in Σ. However, for any interpretation $\star$ in $\mathbf{I\Delta_0} + \mathbf{Exp}$, both $A^\star$ and $\neg A^\star$ are Σ_1, because $A^\star$ is provably equivalent to $0 = 0$, and $\neg A^\star$ is provably equivalent to $0 = 1$. ∎

We conclude the paper with a conjecture:

Conjecture. A formula A of **ILP** is essentially Σ_1 with respect to $\mathbf{I\Delta_0} + \mathbf{Exp}$ iff there is a formula $B \in \Sigma$ such that $\mathbf{L} \vdash A^\circ \leftrightarrow B^\circ$.

References

[DJP] D. De Jongh, Duccio Pianigiani, *Solution of a problem of David Guaspari*, Studia Logica, (1996).

[DV] D. De Jongh, F. Veltman, *Provability logics for relative interpretability*, P. Petkov editor, Mathematical Logic, Proceedings of the Heyting 1988 summer school in Varna, Bulgaria, Plenum Press, Boston 1990.

[G] D. Guaspari, *Sentences implying their own provability*, Journal of Symbolic Logic 48 (1983), 777-789.

[GJ] E. Goris, J.J. Joosten, *Modal matters in interpretability logic*, Technical Report LGPS 226 University of Utrecht 2003.

[GS] D. Guaspari, R. Solovay, *Rosser sentences*, Annals of Mathematical Logic.

[HM] P. Hájek, F. Montagna, *The logic of* Π_1 *conservativity continued*, Archive for Mathematical Logic 32 (1992), 57-63.

[K] M. Kalsbeck, *Towards the interpretability logic of* $I\Delta_0 + Exp$, Technical Report LGPS 61, University of Utrecht 1990.

[Sm] C. Smorynski, *Self-reference and modal logic*, Springer Verlag 1985.

[V1] A. Visser, *Interpretability Logic*, P. Petkov editor, Mathematical Logic, Proceedings of the Heyting 1988 summer school in Varna, Bulgaria, Plenum Press, Boston 1990.

[V2] A. Visser, *An overview of interpretability Logic*, Technical Report 174, University of Utrecht 1997.

[V3] A. Visser, *A course on bimodal provability logic*, Annals of Pure and Applied Logic (1995), 109-142.