Math. Log. Quart. **0** (0) 0, 1 – 4

Mathematical Logic
Quarterly

Shoenfield is Gödel after Krivine

Thomas Streicher[a 1)] **and Ulrich Kohlenbach** [a 2)]

[a] Fachbereich Mathematik, Technische Universität Darmstadt, Schloßgartenstraße 7, 64289, Darmstadt, Germany

Abstract. We show that Shoenfield's functional interpretation of Peano arithmetic can be factorized as a negative translation due to J. L. Krivine followed by Gödel's Dialectica interpretation.

Mathematics Subject Classification: 03F03, 03F10.

Keywords: functional interpretation, negative translation.

1 Introduction

For unwinding classical proofs of Π_2 statements $A \equiv \forall n \exists m R(n, m)$ in Peano arithmetic PA a convenient way is to consider the functional interpretation of some negative translation of A giving rise to a Gödel T functional f for which HA^ω (see [Tr73]) proves $\forall n R(n, f(n))$.

In his book [Sh67] J. Shoenfield introduced a functional interpretation for Peano arithmetic PA associating with every formula A a formula $A^S \equiv \forall u \exists x A_S(u, x)$ (here u and x stand for lists of variables) with A_S quantifier-free by recursion on the structure of A in the following way

(S1) $P^S \equiv P \equiv P_S$ for prime P
(S2) $(\neg A)^S \equiv \forall f \exists u \, \neg A_S(u, f(u))$
(S3) $(A \vee B)^S \equiv \forall uv \exists xy \, A_S(u, x) \vee B_S(v, y)$
(S4) $(\forall z \, A)^S \equiv \forall zu \exists x \, A_S(z, u, x)$

where $A^S \equiv \forall u \exists x A_S(u, x)$ and $B^S \equiv \forall v \exists y B_S(v, y)$. One can show that whenever $\mathsf{PA} \vdash A$ then $\mathsf{HA}^\omega \vdash \forall u A_S(u, t(u))$ for some HA^ω term t.

There arises the question whether for some appropriate negative translation $(-)'$ it holds that Gödel's functional interpretation of A' coincides with $\exists f \forall u A_S(u, f(u))$ for all arithmetic A. In this note we will give a positive answer to this question using a negative translation inspired by J. L. Krivine's [Kr90]. This result was already

1) e-mail: streicher@mathematik.tu-darmstadt.de
2) e-mail: kohlenbach@mathematik.tu-darmstadt.de

announced in a side remark in [Hy02] (p. 63) without giving any details and obtained independently by J. Avigad [Av06].

In section 2 we introduce Krivine's negative translation, in section 3 we prove our main result that Shoenfield's functional interpretation can be obtained as Krivine's negative translation followed by Gödel's functional (Dialectica) interpretation and finally in section 4 we discuss the extension of our main result to the remaining logical connectives $\wedge$ and $\exists$.

2 Krivine' Negative Translation

In [Kr90] J.-L. Krivine introduced a particularly simple negative translation for second order predicate logic (formulated in the $\to\forall$-fragment) which just inserts a negation in front of every prime formula. It has been extended to the remaining logical connectives in [SR98] where one can find also applications to the theory of functional programming (abstract machines, continuations etc.). For the purposes of this note we, however, prefer to introduce Krivine's negative translation as an optimized variant of Kuroda's negative translation (see [Ku51]).

Kuroda's negative translation is defined as $A^K \equiv \neg\neg A^\dagger$ where $(-)^\dagger$ is defined inductively as

(1) $P^\dagger \equiv P$ for prime P
(2) $(A \Box B)^\dagger \equiv A^\dagger \Box B^\dagger$ for $\Box \in \{\wedge, \vee, \to\}$
(3) $(\exists x A)^\dagger \equiv \exists x A^\dagger$
(4) $(\forall x A)^\dagger \equiv \forall x \neg\neg A^\dagger$.

We write A° for $\neg A^\dagger$. For A in the $\neg\vee\forall$-fragment we have

(i) $P^\circ \equiv \neg P^\dagger \equiv \neg P$
(ii) $(\neg A)^\circ \equiv \neg(\neg A)^\dagger \equiv \neg\neg A^\dagger \equiv \neg A^\circ$
(iii) $(A \vee B)^\circ \equiv \neg(A \vee B)^\dagger \equiv \neg(A^\dagger \vee B^\dagger) \iff \neg A^\dagger \wedge \neg B^\dagger \equiv A^\circ \wedge B^\circ$
(iv) $(\forall x A)^\circ \equiv \neg(\forall x A)^\dagger \equiv \neg\forall x \neg\neg A^\dagger \iff \neg\neg\exists x \neg A^\dagger \equiv \neg\neg\exists x A^\circ$

where $\iff$ stands for intuitionistic equivalence. Since $A^K \equiv \neg A^\circ$ and $\neg\neg\forall x \neg\neg A^\dagger \iff \neg\neg\neg\exists x \neg A^\dagger \iff \neg\exists x A^\circ$ the Kuroda negative translation for the $\neg\vee\forall$-fragment is not optimal for $\forall$ since it introduces a triple negation where a single negation would suffice. This observation suggests the following negative translation for the $\neg\vee\forall$-fragment which was considered in [SR98] with a different motivation: $A' \equiv \neg A^*$ where A^* is defined inductively as

(K1) $P^* \equiv \neg P$ if P is prime
(K2) $(\neg A)^* \equiv \neg A^*$
(K3) $(A \vee B)^* \equiv A^* \wedge B^*$
(K4) $(\forall x A)^* \equiv \exists x A^*$.

If $(-)^*$ is replaced by $(-)^\circ$ these are the same as the equivalences given by (i)-(iv), except that a double-negation has been omitted in (K4). Nevertheless we have that

Proposition 2.1. *$A' \iff A^K$ for all A in the $\neg\vee\forall$-fragment.*

Proof. We proceed by induction on the structure of formulas. For the cases of prime formulas and negations the argument is trivial.

For disjunctions $A \vee B$ the claim follows from the equivalences

$$(A \vee B)' \equiv \neg(A^* \wedge B^*) \iff A^* \to \neg B^* \iff \neg\neg A^* \to \neg B^* \iff \neg A' \to B'$$

$$(A \vee B)^K \equiv \neg(A^\circ \wedge B^\circ) \iff A^\circ \to \neg B^\circ \iff \neg\neg A^\circ \to \neg B^\circ \equiv \neg A^K \to B^K$$

together with the induction hypothesis for A and B.

For $\forall x A$ the claim follows from the equivalences

$$(\forall x A)' \equiv \neg\exists x A^* \iff \neg\neg\neg\exists x A^* \iff \neg\neg\forall x \neg A^* \equiv \neg\neg\forall x A'$$

$$(\forall x A)^K \equiv \neg(\forall x A)^\circ \iff \neg\neg\neg\exists x A^\circ \iff \neg\neg\forall x \neg A^\circ \equiv \neg\neg\forall x A^K$$

together with the induction hypothesis for A. □

Thus A' is also intuitionistically equivalent to the Gödel-Gentzen negative translation of A (since this holds for the Kuroda translation of A).

3 Shoenfield is Gödel after Krivine

For convenience we first recall the definition of Gödel's functional interpretation.

(D1) $P^D \equiv P \equiv P_D$ if P is prime

(D2) $(\neg A)^D \equiv \exists f \forall u \neg A_D(u, f(u))$

(D3) $(A \wedge B)^D \equiv \exists uv \forall xy (A_D(u,x) \wedge B_D(v,y))$

(D4) $(\forall z A)^D \equiv \exists f \forall z x A_D(z, f(z), x)$

(D5) $(A \to B)^D \equiv \exists f g \forall u y (A_D(u, g(u,y)) \to B_D(f(u), y))$

(D6) $(\exists z A)^D \equiv \exists z u \forall x A_D(z, u, x)$

(D7) $(A \vee B)^D \equiv \exists n u v \forall x y \ (n = 0 \to A_D(u,x)) \wedge (n \neq 0 \to B_D(v,y))$

Now we can prove our main result.

Theorem 3.1. *For every arithmetic formula A it holds that*

(1) $A^*_D(u,x) \iff \neg A_S(u,x)$ *where* $(A^*)^D \equiv \exists u \forall x\, A^*_D(u,x)$

(2) $A'_D(f,u) \iff A_S(u, f(u))$ *where* $(A')^D \equiv \exists f \forall u\, A'_D(f,u)$

where $\iff$ *stands for provably equivalent in* HA^ω.

Proof. First we show that for every formula A condition (1) implies condition (2). We have $(A')^D \equiv \exists f \forall u\, A'_D(f,u)$ with $A'_D(f,u) \equiv \neg A^*_D(u, f(u))$. From (1) we know $A^*_D(u, f(u)) \iff \neg A_S(u, f(u))$ and, accordingly, we have

$$A'_D(f,u) \iff \neg A^*_D(u, f(u)) \iff \neg\neg A_S(u, f(u)) \iff A_S(u, f(u))$$

as desired.

Next we prove (1) by induction on the structure of A. The base case is trivial.

We have $((\neg A)^*)^D \equiv \exists f \forall u\, ((\neg A)^*)_D(f,u)$ where $((\neg A)^*)_D(f,u) \equiv (\neg A^*)_D\ (f,u)$ $\equiv \neg A^*_D(u, f(u))$. By induction hypothesis we have $A^*_D(u, f(u)) \iff \neg A_S(u, f(u))$ and thus $((\neg A)^*)_D(f,u) \iff \neg\neg A_S(u, f(u)) \equiv \neg(\neg A)_S(f,u)$ as desired.

We have $((A \vee B)^*)^D \equiv \exists uv \forall xy\, (A \vee B)^*_D(u,v,x,y)$ with $(A \vee B)^*_D\ (u,v,x,y) \equiv$ $(A^* \wedge B^*)_D(u,v,x,y) \equiv A^*_D(u,x) \wedge B^*_D(v,y)$. By induction hypothesis we have $A^*_D(u,x) \iff \neg A_S(u,x)$ and $B^*_D(v,y) \iff \neg B_S(v,y)$ from which it follows that $(A \vee B)^*_D(u,v,x,y) \iff \neg A_S(u,x) \wedge \neg B_S(v,y) \iff \neg(A_S(u,x) \vee B_S(v,y)) \equiv$ $\neg(A \vee B)_S(u,v,x,y)$ as desired.

We have $((\forall z A)^*)^D \equiv \exists zu \forall x\, (\forall z A)^*_D(z,u,x)$ with $(\forall z A)^*_D \equiv (\exists z A^*)_D \equiv A^*_D$. By induction hypothesis we have $A^*_D(z,u,x) \iff \neg(A^*)_S(z,u,x)$ and thus obtain $(\forall z A)^*_D(z,u,x) \iff \neg(A^*)_S(z,u,x) \equiv \neg(\forall z A^*)_S(z,u,x)$ as desired. $\square$

4 Extension to $\wedge$ and $\exists$

As usual in classical logic one defines $A \to B \equiv \neg A \vee B$ and $\exists x\, A(x) \equiv \neg\forall x \neg A(x)$. Thus, we get as derived clauses for $(-)^*$

(K5) $(A \to B)^* \equiv A' \wedge B^*$ and thus $(A \to B)' \iff A' \to B'$

(K6) $(\exists x A(x))^* \equiv \neg\exists x \neg A^*(x)$ and thus
$(\exists x A(x))' \equiv \neg\neg\exists x \neg A^*(x) \equiv \neg\neg\exists x A' \iff \neg\forall x \neg A'$.

Although one could define $A \wedge B$ as $\neg(\neg A \vee \neg B)$ it turns out as simpler to define $(-)^*$ for conjunction directly as

(K7) $(A \wedge B)^* \equiv A^* \vee B^*$ and thus $(A \wedge B)' \iff A' \wedge B'$.

Thus, in order to keep Theorem 3.1 valid we extend Shoenfield's functional interpretation of PA to the remaining connectives as follows

(S5) $(A{\to}B)^S \equiv \forall f v \exists u y\, A_S(u, f(u)) \to B_S(v,y)$

(S6) $(\exists z A)^S \equiv \forall U \exists z f\, A_S(z, U(z,f), f(U(z,f)))$

(S7) $(A \wedge B)^S \equiv \forall n u v \exists x y\, (n{=}0 \to A_S(u,x)) \wedge (n{\neq}0 \to B_S(v,y))$.

Notice that (S6) is obtained from $(\forall z \neg A)^S \equiv \forall z f \exists u \neg A_S(z,u,f(u))$ by applying (S2). Admittedly, the clause (S6) does not look very nice but if A is quantifierfree we get $(\exists z A)^S \equiv \exists z A$ since u and x are empty lists of variables.

The somewhat strange form of (S7) is enforced by the Dialectica interpretation of disjunction since we have

$$\begin{aligned}
(A \wedge B)^*_D(n,u,v,x,y) &\equiv (A^* \vee B^*)_D(n,u,v,x,y) \\
&\equiv (n{=}0 \to A^*_D(u,x)) \wedge (n{\neq}0 \to B^*_D(v,y)) \\
&\iff (n{=}0 \to \neg A^*_S(u,x)) \wedge (n{\neq}0 \to \neg B^*_S(v,y)) \\
&\iff \neg((n{=}0 \wedge A^*_S(u,x)) \vee ((n{\neq}0 \wedge B^*_S(v,y))) \\
&\iff \neg((n{=}0 \to A^*_S(u,x)) \wedge ((n{\neq}0 \to B^*_S(v,y))) \\
&\equiv \neg(A \wedge B)_S(n,u,v,x,y)
\end{aligned}$$

and thus $(A \wedge B)_S(n,u,v,x,y) \equiv (n{=}0 \to A^*_S(u,x)) \wedge (n{\neq}0 \to B^*_S(v,y))$. Notice, however, that $(A \wedge B)^S \iff \forall uv \exists xy A_S(u,x) \wedge B_S(v,y)$ as one might expect.

References

[Av06] J. Avigad *A variant of the double negation translation.* Carnegie Mellon Technical Report CMU-PHIL 179 (2006).

[Hy02] J. M. E. Hyland *Proof theory in the abstract.* Ann. Pure Appl. Logic 114, no. 1-3, pp.43–78 (2002).

[Kr90] J.-L. Krivine *Opérateurs de mise en mémoire et traduction de Gödel.* Arch. Math. Logic 30, no.4, pp.241-267 (1990).

[Ku51] S. Kuroda *Intuitionistische Untersuchungen der formalistischen Logik.* Nagoya Math. 3, ppp. 35-47 (1951).

[Sh67] J. Shoenfield *Mathematical Logic.* Addison-Wesley Publishing Co. (1967).

[SR98] T. Streicher and B. Reus *Classical logic, continuation semantics and abstract machines.* J. Funct. Prog. 8(6), pp.543-572 (1998).

[Tr73] A. Troelstra (ed.) *Metamathematical Investigations of Intuitionistic Arithmetic and Analysis.* SLNM 344, Springer Verlag, 1973.