Shoenfield = Gödel ∘ Krivine

T. Streicher and U. Kohlenbach

August 23, 2006

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 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 $A^S \equiv \forall u \exists x A_S(u,x)$ with A_S quantifier-free by recursion on the structure of A in the following way

(S1) $P^S \equiv P \equiv P_S$

(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 $B^S \equiv \forall v \exists y B_S(v,y)$.

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. For this purpose one can use a negative translation studied by T. Streicher and B. Reus in [SR98]. It is inspired by a surprisingly simple negative translation due to J.-L. Krivine for second order predicate logic (formulated in the $\to\forall$-fragment) which just inserts a negation in front of every prime formula. For our purpose we prefer to introduce this negative translation as an optimized variant of Kuroda's negative translation.

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 $\neg\vee\exists\forall$-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 logical 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 transaltion 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) $(\forall x\, A)^* \equiv \exists x\, A^*$.

The next theorem relates the Shoenfield translation to Gödel's functional interpretation of our negative translation. For convenience we recall Gödel's functional interpretation.

(D1) $P^D \equiv P$

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

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

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

Theorem 0.1 *For every arithemtic 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 (1) implies (2). We have $(A')^D(f, u) \equiv \exists f \forall u\, A'_D(f, u)$ 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 $(\forall z A)^*_D(z,u,x) \iff \neg(A^*)_S(z,u,x) \equiv \neg(\forall z A^*)_S(z,u,x)$ as desired. □

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 0.1 valid we extend Shoenfield's functional interpretation of PA to the remaining connectives as follows

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

(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 quantifier-free 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 imposed on us by the Dialectica interpretation of disjunction since we have $(A \wedge B)^*_D(n,u,v,x,y) \equiv (A^* \vee B^*)_D(n,u,v,x,y) \equiv (n{=}0 \to A^*_D(u,x)) \wedge (n{\neq}0 \to B^*_D(v,y)) \iff (n{=}0 \to \neg A^*_S(u,x)) \wedge (n{\neq}0 \to \neg B^*_S(v,y)) \iff \neg((n{=}0 \wedge A^*_S(u,x)) \vee ((n{\neq}0 \wedge B^*_S(v,y))) \iff \neg((n{=}0 \to A^*_S(u,x)) \wedge ((n{\neq}0 \to B^*_S(v,y))) \equiv \neg(A \wedge B)_S(n,u,v,x,y)$ 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 $(A \wedge B)^S \iff \forall uv \exists xy A_S(u,x) \wedge B_S(y,y)$ as one might expect.

References

[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).