Believability theories

Corrigendum to: *The Epistemic Lightness of Truth. Deflationism and its Logic*

Cezary Cieśliński

In (Cieśliński, 2017, p. 254) the following definition of a believability theory over K is presented:

Definition 1 *Let K be an axiomatisable extension of PA in the language L_K (which is possibly richer than L_{PA}). Denote as $L_{K,B}$ the extension of L_K with a new one-place predicate 'B'. Let KB be a theory K formulated in the language $L_{K,B}$.*

- *We denote as $Bel(K)^-$ the theory in the language $L_{K,B}$ which extends KB with the following axioms:*

  (A_1) $\forall\psi \in L_{K,B}[KB \vdash \psi \rightarrow B(\psi)]$

  (A_2) $\forall\varphi, \psi \in L_{K,B}[\big(B(\varphi) \wedge B(\varphi \rightarrow \psi)\big) \rightarrow B(\psi)]$

  In addition, the theory $Bel(K)^-$ has the following rules of inference:

$$\text{NEC}\quad \frac{\vdash \phi}{\vdash B(\phi)} \qquad\qquad \frac{\vdash \forall x B\phi(x)}{\vdash B\big(\forall x\phi(x)\big)}\quad \text{GEN}$$

- *We denote as $Bel(K)$ a theory which is exactly like $Bel(K)^-$, except that it contains all the axioms of induction for formulas of $L_{K,B}$.*

In general, results about believability of various statements depend on the choice of K. Later in the book it is claimed that $Bel(PA)$ permits us to prove the believability of reflection principles for Peano arithmetic and $Bel(TB^-)$ proves the believability of all the axioms of CT (the classical compositional truth theory with full extended induction).

In fact, however, both results require iterating believability. This operation is defined in the first section of this note, where (for illustration) I derive also the believability of uniform reflection for Peano arithmetic. Similarly, deriving the believability of all the axioms of CT can be achieved in this framework by iterating believability over TB^-.

In the second section of this Corrigendum I present a slightly modified approach to believability which is the one I currently favour. Namely, I introduce the theory $Bel^*(K)$, in which the GEN rule is replaced with an appropriate axiom. Then I show how to obtain both the believability of reflection and the believability of CT directly in $Bel^*(K)$, without any need of further iterations.

1 Iterating believability

Definition 2

- *Let* $Bel_0(PA)^-$ *be* $Bel(PA)^-$,
- *Let* $Bel_{n+1}(PA)^-$ *be the theory which is exactly like* $Bel(PA)^-$, *except that its first axiom has the following form:*

  (A_1) $\forall\psi \in L_{PAB}\big[\big(Bel_n(PA)^- \vdash B(\psi)\big) \to B(\psi)\big]$

In effect, in each successor step we declare that the internal theory *of the theory previously obtained* is believable.

We are going to show that the believability of the uniform reflection for PA is provable after the first iteration is made; in other words, it is provable in $Bel_1(PA)^-$.

Theorem 3 $Bel_1(PA)^- \vdash \forall\varphi(x) \in L_{PA}\ B\big(\forall x[Pr_{PA}(\varphi(x)) \to \varphi(x)]\big)$.

Proof. Working in $Bel_1(PA)^-$, we note that:

$$(*)\ \forall\varphi(x) \in L_{PA}\ Bel_0(PA)^- \vdash B\big(\forall x\forall d\ B(Prov_{PA}(d, \varphi(x)) \to \varphi(x))\big).$$

The reason is (we observe it inside $Bel_1(PA)$!) that given $\varphi(x) \in L_{PA}$, we can argue as follows in $Bel_0(PA)^-$:

- $\forall x, d\ PA \vdash Prov_{PA}(d, \varphi(x)) \to \varphi(x)$,[1]

[1]Cf. the argument given on p. 263 of (Cieśliński, 2017).

- $\forall x, d\ B\big(Prov_{PA}(d, \varphi(x)) \to \varphi(x)\big)$ (by the axiom (A_1) of $Bel_0(PA)$),
- hence (*) is obtained by NEC rule of $Bel_0(PA)$ applied to the previous line.

From (*) we conclude that:

(**) $\forall \varphi(x) \in L_{PA}\ Bel_0(PA)^- \vdash B\big(\forall x \forall d\ (Prov_{PA}(d, \varphi(x)) \to \varphi(x))\big)$.

Observe that (**) is the result of applying the rule GEN of $Bel_0(PA)^-$ to (*). It follows that:

(***) $\forall \varphi(x) \in L_{PA}\ Bel_0(PA)^- \vdash B\big(\forall x[Pr_{PA}(\varphi(x)) \to \varphi(x)]\big)$.

From (***) by the axiom (A_1) of $Bel_1(PA)^-$ we obtain the believability of the uniform reflection principle for Peano arithmetic: $\forall \varphi(x) \in L_{PA}\ B\big(\forall x[Pr_{PA}(\varphi(x)) \to \varphi(x)]\big)$. $\square$

2 Believability without the generalisation rule

In this section the theory $Bel^*(K)$ is introduced, which (in my opinion) is better than $Bel(K)$ as a formal description of the behaviour of the believability predicate. As before, we assume that K is an axiomatisable extension of PA in the language L_K. The expressions '$L_{K,B}$' and 'KB' also retain their previous meanings.

Definition 4 *We denote as $Bel^*(K)$ the theory in the language $L_{K,B}$ which extends KB with the following axioms:*

(A_1) $\forall \psi \in L_{K,B}[KB \vdash \psi \to B(\psi)]$

(A_2) $\forall \varphi, \psi \in L_{K,B}[\big(B(\varphi) \wedge B(\varphi \to \psi)\big) \to B(\psi)]$

(A_3) $\forall \varphi(x) \in L_{K,B}\Big(B\big(\forall x B\varphi(x)\big) \to B\big(\forall x \varphi(x)\big)\Big)$.

In addition, the theory $Bel^*(K)$ has the following rule of inference:

$$\text{NEC} \quad \frac{\vdash \psi}{\vdash B(\psi)}$$

The difference with $Bel(K)$ is that now axiom (A_3) replaces the GEN rule. Indeed, the intuitive motivation for the axiom is exactly the same as the motivation for the rule described on p. 255 of (Cieśliński, 2017).

For $Bel^*(K)$, we obtain the following counterpart of Theorem 13.4.3 (see (Cieśliński, 2017, p. 257)):

Theorem 5 *If N is expandable to a model N^* of K, then N^* is expandable to a model of $Int_{Bel^*_{Con}(K)}$.*

The proof (due to Mateusz Łełyk) consists in adapting my proof of Theorem 13.4.3. As in the book, the expansion (N^*, B) satisfying $Int_{Bel^*_{Con}(K)}$ can be obtained as a supervaluational model. We adopt the following definition.

Definition 6 *For a model N^* of K, we define:*

- $B_0 = KB$,
- $B_{n+1} = \{\psi : \forall Z \supseteq B_n[\mathit{if}\ (N^*, Z) \models (A_1) - (A_3), (\text{CON}),\ \mathit{then}\ (N^*, Z) \models \psi]\}$,
- $B_\omega = \bigcup_{n \in N} B_n$.

Then $(N^*, B_\omega) \models Int_{Bel^*_{Con}(K)}$. The successor steps permit us to handle the NEC rule.

2.1 Application: reflection principles

The first observation is that $Bel^*(PA)$ is quite efficient in proving, within the scope of of 'B', strong reflection principles. Let me start with the following definition of a sequence of stronger and stronger theories, obtained by iterating the uniform reflection principle.

Definition 7

- $S_0 = PAB$
- $S_{n+1} = S_n \cup \{\forall x[Pr_{S_n}(\varphi(x)) \to \varphi(x)] : \varphi(x) \in L_{PAB}\}$.

In other words, we start with PAB (which is PA in the language with the new predicate 'B'), and then, at each stage, we add the uniform reflection principle for the preceding theory, for formulas of the language of PAB.

It turns out that $Bel^*(PA)$ proves the believability of each theory S_n. Accordingly, $Bel^*(PA)$ proves the believability of uniform reflection not just for PA, but for each S_n.

Theorem 8 *For every natural number n, $Bel^*(PA) \vdash \forall\varphi(x) \in L_{PAB}\ B\big(\forall x[Pr_{S_n}(\varphi(x)) \to \varphi(x)]\big)$.*

Proof. For $n = 0$, we argue as follows in $Bel^*(PA)$:

1. $\forall\varphi(x) \in L_{PAB}\ \forall x \forall d\ PAB \vdash Prov_{PAB}(d, \varphi(x)) \to \varphi(x)$ (provable in PAB),
2. $\forall\varphi(x) \in L_{PAB}\ \forall x \forall d\ B\big(Prov_{PAB}(d, \varphi(x)) \to \varphi(x)\big)$ (by (A_1)),
3. $B\Big(\forall\varphi(x) \in L_{PAB}\ \forall x \forall d\ B\big(Prov_{PAB}(d, \varphi(x)) \to \varphi(x)\big)\Big)$ (by NEC)
4. $\forall\varphi(x) \in L_{PAB}\ B\Big(\forall x \forall d\ B\big(Prov_{PAB}(d, \varphi(x)) \to \varphi(x)\big)\Big)$ (by (A_1), (A_2)),
5. $\forall\varphi(x) \in L_{PAB}\ B\Big(\forall x \forall d\ \big(Prov_{PAB}(d, \varphi(x)) \to \varphi(x)\big)\Big)$ (by (A_3)),
6. $\forall\varphi(x) \in L_{PAB}\ B\Big(\forall x\big(Pr_{PAB}(\varphi(x)) \to \varphi(x)\big)\Big)$ (by (A_1), (A_2)).

For the inductive part, assume that for a fixed n:

$$Bel^*(PA) \vdash \forall\varphi(x) \in L_{PAB}\ B\big(\forall x[Pr_{S_n}(\varphi(x)) \to \varphi(x)]\big).$$

We claim that

$$Bel^*(PA) \vdash \forall\varphi(x) \in L_{PAB}\ B\big(\forall x[Pr_{S_{n+1}}(\varphi(x)) \to \varphi(x)]\big).$$

Working in $Bel^*(PA)$, it is enough to demonstrate that:[2]

$$\forall\varphi(x) \in L_{PAB}\ \forall x \forall d\ B\big(Prov_{S_{n+1}}(d, \varphi(x)) \to \varphi(x)\big).$$

Fixing x, d and $\varphi(x) \in L_{PAB}$, we consider cases.

Case 1: If $\neg Prov_{S_{n+1}}(d, \varphi(x))$, then already PA proves that this is so and thus $B\big(\neg Prov_{S_{n+1}}(d, \varphi(x))\big)$, which by (A_1) and (A_2) yields the desired conclusion that $B\big(Prov_{S_{n+1}}(d, \varphi(x)) \to \varphi(x)\big)$.

Case 2: If $Prov_{S_{n+1}}(d, \varphi(x))$, then the following two observations lead to the conclusion that $B\big(\varphi(x)\big)$:

(a) Every axiom of S_{n+1} used in d is believable.

(b) Believability is preserved under Modus Ponens.

For (a), note that by the inductive assumption we have:

$$\forall\psi \in L_{PAB} Pr_{S_n}(\psi) \to B(\psi),^3$$

thus if ψ is an axiom of S_{n+1} which is also an axiom of S_n, we have $B(\psi)$. On the other hand, if ψ is an axiom of S_{n+1} but not of S_n, then it must have the form '$\forall x[Pr_{S_n}(\varphi(x)) \to \varphi(x)]$' and we have $B(\psi)$ by the inductive assumption.

In this way we conclude that $B\big(\varphi(x)\big)$. Therefore also in Case 2 we obtain $B\big(Prov_{S_{n+1}}(d, \varphi(x)) \to \varphi(x)\big)$, as required. $\square$

2.2 Application: recovering compositional axioms from disquotational principles

Let us start with the following fact ('$x \in Var$' reads 'x is a variable').

Fact 9 $Bel^*(TB^-) \vdash \forall\varphi(x) \in L_{PA} \forall a \in Var B\big(\forall a(T(\varphi(a)) \equiv \varphi(a))\big)$.

[2] After this, the proof proceeds exactly as in the step for $n = 0$.

[3] Assume that $Pr_{S_n}(\psi)$. Then $PA \vdash Pr_{S_n}(\psi)$ and so $B(Pr_{S_n}(\psi))$. By the inductive assumption we have: $B(Pr_{S_n}(\psi) \to \psi)$, so by (A_2) we obtain $B(\psi)$.

Proof. We reason in $Bel^*(TB^-)$ as follows:

(1) $\forall\varphi(x) \in L_{PA} \forall x\ TB^- \vdash T(\varphi(x)) \equiv \varphi(x)$

(2) $\forall\varphi(x) \in L_{PA} \forall x\ B\big(T(\varphi(x)) \equiv \varphi(x)\big)$

(3) $B\Big(\forall\varphi(x) \in L_{PA} \forall x\ B\big(T(\varphi(x)) \equiv \varphi(x)\big)\Big)$

(4) $\forall\varphi(x) \in L_{PA} B\Big(\forall x B\big(T(\varphi(x)) \equiv \varphi(x)\big)\Big)$

(5) $\forall\varphi(x) \in L_{PA} B\big(\forall x (T(\varphi(x)) \equiv \varphi(x))\big)$.

(6) $\forall\varphi(x) \in L_{PA} \forall a \in Var B\big(\forall a (T(\varphi(a)) \equiv \varphi(a))\big)$

(1) is provable already in PA; (2) follows from (1) by axiom (A_1); (3) is obtained from (2) by NEC; (5) is obtained from (4) by (A_3), (6) follows from (5) by (A_1) and (A_2) (renaming of variables produces logically equivalent statements).

For (4) we argue as follows: fix $\varphi(x) \in L_{PA}$, then by Σ_1 completeness $PA \vdash \varphi(x) \in L_{PA}$, hence $B(\varphi(x) \in L_{PA})$ by (A_1). Denote as $F\big(\varphi(x)\big)$ the following sentence:

$$\varphi(x) \in L_{PA} \to \Big(\forall\varphi(x) \in L_{PA} \forall x\ B\big(T(\varphi(x)) \equiv \varphi(x)\big) \to \forall x B\big(T(\varphi(x)) \equiv \varphi(x)\big)\Big).$$

Observe that $\emptyset \vdash F\big(\varphi(x)\big)$; therefore by (A_1) we have $B\Big(F\big(\varphi(x)\big)\Big)$ and hence by two applications of (A_2) we obtain $B\Big(\forall x B\big(T(\varphi(x)) \equiv \varphi(x)\big)\Big)$. $\square$

With the fact at hand, we can demonstrate that $Bel^*(TB^-)$ proves the believability of compositional principles of truth.

Theorem 10 $Bel^*(TB^-) \vdash B(CT)$.

The expression '$B(CT)$' is a shorthand of 'truth-theoretic axioms of CT are believable and for every $\varphi \in L_T$ induction for φ is believable'.

Proof. We start by showing that ('$Tm(t)$' reads 't is a constant term'):

$$Bel^*(TB^-) \vdash B\Big(\forall t, s\big[Tm(t) \wedge Tm(s) \rightarrow (T(t = s) \equiv val(t) = val(s))\big]\Big).$$

The reasoning (carried out in $Bel^*(TB^-)$) goes as follows:

(1)	$\forall t, s\ TB^- \vdash Tm(t) \wedge Tm(s) \rightarrow (T(t = s) \equiv val(t) = val(s))$	(provable in PA)
(2)	$\forall t, s\ B\big(Tm(t) \wedge Tm(s) \rightarrow (T(t = s) \equiv val(t) = val(s))\big)$	(axiom (A_1))
(3)	$B\Big(\forall t, s\ B\big(Tm(t) \wedge Tm(s) \rightarrow (T(t = s) \equiv val(t) = val(s))\big)\Big)$	(by NEC)
(4)	$B\Big(\forall t, s\big[Tm(t) \wedge Tm(s) \rightarrow (T(t = s) \equiv val(t) = val(s))\big]\Big)$	(by (A_3))

As for compositional axioms for sentential connectives, only the case of negation will be considered here. We claim that:

$$Bel^*(TB^-) \vdash B\big(\forall \psi[T(\neg\psi) \equiv \neg T(\psi)]\big).$$

The reasoning (carried out in $Bel^*(TB^-)$) goes as follows:

(1)	$\forall \psi\ TB^- \vdash T(\neg\psi) \equiv \neg T(\psi)$	(provable in PA)
(2)	$\forall \psi\ B\big(T(\neg\psi) \equiv \neg T(\psi)\big)$	(axiom (A_1))
(3)	$B\Big(\forall \psi\ B\big(T(\neg\psi) \equiv \neg T(\psi)\big)\Big)$	(by NEC)
(3)	$B\big(\forall \psi[T(\neg\psi) \equiv \neg T(\psi)]\big)$	(by (A_3))

We now consider the case of the existential quantifier. It will be demonstrated that:

$$Bel^*(TB^-) \vdash B\Big(\forall \varphi(x) \in L_{PA} \forall a \in Var\big[T(\exists a \varphi) \equiv \exists x T(\varphi(x))\big]\Big).$$

Working in $Bel^*(TB^-)$, we reason as follows:

(1)	$\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(\forall a(T(\varphi(a)) \equiv \varphi(a))\big)$	(Fact 9)
(2)	$\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(\exists a T(\varphi(a)) \equiv \exists a \varphi(a)\big)$	((A_1) and (A_2), (1))
(3)	$\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(T(\exists a \varphi(a)) \equiv \exists a \varphi(a)\big)$	(A_1)
(4)	$\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(T(\exists a \varphi(a)) \equiv \exists a T(\varphi(a))\big)$	(from (2) and (3))
(5)	$\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(T(\exists a \varphi(a)) \equiv \exists x T(\varphi(x))\big)$	(variable renaming)
(6)	$B\Big(\forall \varphi(x) \in L_{PA} \forall a \in Var\ B\big(T(\exists a \varphi(a)) \equiv \exists x T(\varphi(x))\big)\Big)$	(NEC)
(7)	$B\Big(\forall \varphi(x) \in L_{PA} \forall a \in Var\ \big[T(\exists a \varphi(a)) \equiv \exists x T(\varphi(x))\big]\Big)$	(by (A_3))

In the last part of the proof, we check the believability of the extended induction (for the language L_T with the truth predicate). We demonstrate that:

$$Bel^*(TB^-) \vdash \forall\varphi(x) \in L_T B\Big(\big[\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(x+1))\big] \to \forall x\varphi(x)\Big).$$

We reason in $Bel^*(TB^-)$ as follows:

(1) $\forall\varphi(x) \in L_T\ \forall x\ TB^- \vdash \big[\varphi(0) \wedge \forall x\big(\varphi(x) \to \varphi(x+1)\big)\big] \to \varphi(x)$

(2) $\forall\varphi(x) \in L_T\ \forall x\ B\Big(\big[\varphi(0) \wedge \forall x\big(\varphi(x) \to \varphi(x+1)\big)\big] \to \varphi(x)\Big)$

(3) $B\Big(\forall\varphi(x) \in L_T\ \forall x\ B\big(\big[\varphi(0) \wedge \forall x\big(\varphi(x) \to \varphi(x+1)\big)\big] \to \varphi(x)\big)\Big)$

(4) $\forall\varphi(x) \in L_T\ B\Big(\forall x\ B\big(\big[\varphi(0) \wedge \forall x\big(\varphi(x) \to \varphi(x+1)\big)\big] \to \varphi(x)\big)\Big)$

(5) $\forall\varphi(x) \in L_T B\Big(\big[\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(x+1))\big] \to \forall x\varphi(x)\Big).$

For (1), observe that for every $\varphi(x) \in L_T$ and for every x, the formula in question is provable already in pure logic (hence also in TB^-). (2) is obtained from (1) by (A_1), in (3) the NEC rule is applied. For (4), see the justification of a very similar step (4) in the proof of Fact 9. Finally, (5) is obtained from (4) by (A_3). □

References

Cezary Cieśliński. *The Epistemic Lightness of Truth. Deflationism and its Logic.* Cambridge University Press, 2017.