Reflection Principles and Nonclassical Truth

Carlo Nicolai

KING'S College LONDON

Understanding Reflection, München, June 2022
(slides at carlonicolai.github.io)

'Literally speaking, the intended reflection principle cannot be formulated in T itself by means of a single statement. This would require a truth definition...' (Kreisel and Levy 1968, p. 98)

If T is taken to be PA, the compositional, Tarskian theory CT, obtained by turning the definitional clauses of 'truth in $\mathbb{N}$' into axioms suffices for this as it proves its Global Reflection Principle ($\mathrm{GRF}(T)$), and therefore its Uniform Reflection Principle ($\mathrm{RFN}(T)$).

CT can be iterated, and its autonomous iteration $\mathrm{RT}_{<\Gamma_0}$ will then include the autonoumous progression of Uniform Reflection for PA (the limit of such a progression is $\varphi_2 0$). KF (and its schematic extensions) are elegant ways of recapturing these iterations by means of a single, type-free truth predicate.

The interplay of disquotation principles "A' is true iff A' and Uniform Reflection enables one to derive stronger principles for truth: mostly *compositional principles* in fully general form – i.e. with quantification over potentially nonstandard formulae – and *transfinite induction principles*.

In classical logic, an elegant characterization of the interplay between truth and reflection is given by (assuming a canonical notation for ordinals $< \Gamma_0$):

Theorem (Leigh)

- *ε_α induction together with Tarskian truth yields an identical theory as α iterations of Uniform Reflection over typed (uniform) disquotation;*
- *ε_α induction together with Kripke-Feferman truth yields an identical theory as α iterations of Uniform Reflection over type-free, positive (uniform) disquotation.*

In classical logic, disquotation cannot be *full*, while Global and Uniform Reflection Principles are *provably* distinct.

Consider FDE or its standard three-valued extensions K3, LP, S3. Formulate PA over these logics and add (with $\Rightarrow$ a metalinguistic sequent arrow):

$$\mathrm{Tr}\ulcorner A(\overline{x})\urcorner \Leftrightarrow A(x) \tag{$\mathrm{UTB}^{\Rightarrow}$}$$

Lemma

Over $T \supseteq \mathrm{UTB}^{\Rightarrow}$, the following are equivalent:

$$\mathrm{Prov}_T(\varphi) \Rightarrow \mathrm{Tr}\,\varphi \qquad\qquad \mathrm{Prov}_T(\ulcorner A(\overline{x})\urcorner) \Rightarrow A(x)$$

Despite the presence of full disquotation, one lacks a decent conditional.

To obtain significant extensions (including compositionality and transfinite induction), one requires Uniform Reflection for *admissible rules*:

$$\frac{\Rightarrow \mathrm{Prv}_{\mathrm{S}}(\ulcorner \Gamma[\overline{x}] \Rightarrow \Delta[\overline{x}]\urcorner, \ulcorner \Theta[\overline{x}] \Rightarrow \Lambda[\overline{x}]\urcorner) \qquad \Gamma[x] \Rightarrow \Delta[x]}{\Theta[x] \Rightarrow \Lambda[x]} \ (\mathrm{RR}(S))$$

Proposition

$\mathrm{RR}^{\omega}(\mathrm{UTB}^{\Rightarrow})$ *proves all instances of transfinite induction for* $\mathcal{L}_{\mathrm{Tr}}$ *up to* ω^{ω^2}*, as well as all compositional principles for* $\mathcal{L}_{\mathrm{Tr}}$ *(i.e. it also includes the theory* PKF*).*

We can extend FDE with an *intuitionistic* conditional (Leitgeb, Odintsov, Wansing).

Proof-theoretically, one can add rules for the conditional to FDE:

$$\frac{\Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{A \to B, \Gamma \Rightarrow \Delta} \qquad\qquad \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \to B}$$

Semantically, one considers a *Routley Frame* $\mathcal{F} = (S, \leq, \star)$, with $S \neq \varnothing$, $\leq$ a preorder, $\star$ antimonotone and involutive.

Constant domain models $\mathcal{M} = (\mathcal{F}, D, \mathcal{I})$ are defined in the usual way, with $s_0 \leq s_1$ only if $\mathcal{I}_{s_0}(P) \subseteq \mathcal{I}_{s_1}(P)$ and

$$\mathcal{M}, s \vDash \neg A \text{ iff } \mathcal{M}, s^\star \nvDash A$$
$$\mathcal{M}, s \vDash A \to B \text{ iff for } s' \geq s : \mathcal{M}, s' \vDash A \text{ only if } \mathcal{M}, s' \vDash B.$$

Unlike PA formulated in Kleene logics, PA over HYPE can establish the *progressiveness* of the Gentzen jump formula for A

$$A'(z) :\leftrightarrow \forall x \in \mathcal{O}\big(\forall y \prec x\, A(x) \to \forall y \prec x + \omega^z\, A(y)\big)$$

with $\mathcal{O}$ a PR representation system for ordinals $\alpha < \Gamma_0$.

Lemma (Fischer, N., Dopico, 2021)

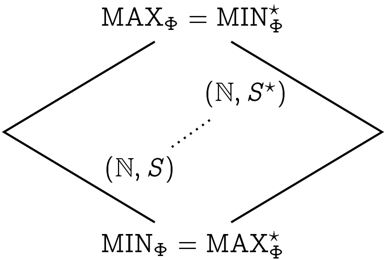
PA *in* HYPE *proves transfinite induction up to any* $\alpha < \varepsilon_0$*.*

The proof of this (and some results below) consists in carefully verifying that the restricted conditional introduction rule suffices to carry out Gentzen's standard argument.

HYPE with Truth

With $\Phi(\cdot)\colon \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ associated with the '4-valued' Kleene evaluation schema, we can generate a Routley Frame by:

$$\begin{aligned}
&\mathbb{S} = \{(\mathbb{N}, X) \mid X \subseteq \mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}} \;\&\; \Phi(X) = X\}\\
&X_0 \leq X_1 :\leftrightarrow X_0 \subseteq X_1\\
&X^{\star} = \mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}} \setminus \{\neg A \mid A \in X\}
\end{aligned}$$



The following principles in $\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}$ are *sound* wrt the model just given (here $\varphi, \psi \in \mathcal{L}_{\mathrm{Tr}}$, i.e. *not* including $\rightarrow$), we call them KFL:

$$\mathrm{Tr}\,(s = t) \leftrightarrow \mathrm{val}(s) = \mathrm{val}(t) \qquad \mathrm{Tr}\,\mathrm{Tr}\,t \leftrightarrow \mathrm{Tr}\,\mathrm{val}(t)$$

$$\mathrm{Tr}\,\neg\varphi \leftrightarrow \neg\mathrm{Tr}\,\varphi \qquad \mathrm{Tr}\,(\varphi \wedge \psi) \leftrightarrow (\mathrm{Tr}\,\varphi \wedge \mathrm{Tr}\,\psi)$$

$$\mathrm{Tr}\,(\forall v \varphi) \leftrightarrow \forall x\,\mathrm{Tr}\,\varphi(\overline{x}/v)$$

Lemma

KFL $\vdash \forall x(\mathrm{Tr}\,\ulcorner A(\overline{x})\urcorner \leftrightarrow A(x))$ *for all* $A(v) \in \mathcal{L}_{\mathrm{Tr}}$ *.*

KFL is remarkably strong, while being able to obtain the T-sentences for $\mathcal{L}_{\mathrm{Tr}}$ in the object language.

Proposition (Fischer, N., Dopico)

KFL *defines hierarchies of Tarskian truth up to any* $\alpha < \varepsilon_0$*. Moreover, its truth predicate is definable in* KF*. Therefore,* KFL $\equiv_{\mathcal{L}_{\mathbb{N}}}$ KF $\equiv_{\mathcal{L}_{\mathbb{N}}}$ $\mathrm{ACA}^{<\varepsilon_0} \equiv_{\mathcal{L}_{\mathbb{N}}} \mathrm{PA} + \mathrm{TI}_{\mathcal{L}_{\mathbb{N}}}(< \varphi_{\varepsilon_0}0)$.

Let $\mathrm{UTB}_0^{\to}$ be the extension of EA in HYPE with the schema $\mathrm{Tr}\ulcorner A(\overline{x})\urcorner \leftrightarrow A(x)$ for $A(v) \in \mathcal{L}_{\mathrm{Tr}}$.

Lemma

Over $T \supseteq \mathrm{UTB}_0^{\to}$*, the following are equivalent:*

1. $(\forall \varphi \in \mathcal{L}_{\mathrm{Tr}})(\mathrm{Prov}_T(\varphi) \to \mathrm{Tr}(\varphi))$ – *or* $\mathrm{Prov}_T(\varphi) \Rightarrow \mathrm{Tr}(\varphi)$
2. $\forall x(\mathrm{Prov}_T(\ulcorner A(\overline{x})\urcorner) \to A(x))$ – *or* $\mathrm{Prov}_T(\ulcorner A(\overline{x})\urcorner) \Rightarrow A(x)$

$(1 \Rightarrow 2)$: immediate (always paying attention to $\to$...)

$(2 \Rightarrow 1)$: we have $\mathrm{Prov}_S(\varphi) \Rightarrow \mathrm{Prov}_S(\mathrm{Tr}\,\overline{\varphi})$, so a cut with the appropriate instance of 2 yields:

$$\mathrm{Prov}_S(\varphi) \Rightarrow \mathrm{Tr}\,\varphi$$

The conditional $\to$ can then be safely introduced.

Proposition

KFL *is a subtheory of* $\mathrm{R}(\mathrm{UTB}_0^{\rightarrow})$, *which is a subtheory of reflecting twice over simple disquotation in* HYPE.

The argument is standard. E.g., for the compositional principles, $\mathrm{EA}^{\rightarrow}$ verifies that, for each $A(v) \in \mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}}$,

$$\mathrm{UTB}_0^{\rightarrow} \vdash \forall x \mathrm{Tr}\,(\ulcorner A(\overline{x})\urcorner) \leftrightarrow \mathrm{Tr}\,(\ulcorner \forall x A \urcorner)$$

The provability of

$$\Rightarrow (\forall \varphi(v) \in \mathcal{L}_{\mathrm{Tr}})(\forall x \mathrm{Tr}\,\varphi(\overline{x}) \leftrightarrow \mathrm{Tr}\,(\forall x \varphi))$$

then follows from Uniform Reflection (to verify this in detail with the restricted conditional intro, one can split the claim in two directions).

Proposition

KFL *is a subtheory of* $\mathrm{R}(\mathrm{UTB}_0^{\rightarrow})$, *which is a subtheory of reflecting twice over simple disquotation in* HYPE.

For full $\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}$-induction, we have, for each $A(v) \in \mathcal{L}_{\mathrm{Tr}}^{\rightarrow}$:

$$\mathrm{EA}^{\rightarrow} \vdash \Rightarrow \mathrm{Prov}_{\mathrm{EA}^{\rightarrow}}(\ulcorner A(0) \wedge \forall x(A(x) \rightarrow A(x+1)) \rightarrow A(\overline{y})\urcorner)$$

Again one application of reflection yields the result.

Starting with the simple bi-conditionals $\mathrm{TB}^{\rightarrow}$, one obtains (as in the classical case) $\mathrm{UTB}^{\rightarrow}$ with one reflection step, and therefore full KFL in two steps, as argued above.

The epsilon numbers ε_α enumerate the fixed points of the function $\alpha \mapsto \omega^\alpha$. We focus on ordinals $< \Gamma_0$. Let:

$$\mathrm{KFL}_{\varepsilon_\alpha} = \mathrm{KFL} + \mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(< \varepsilon_\alpha)$$

with $\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(< \varepsilon_\alpha)$ being

$$\{\mathrm{Prog}(A) \rightarrow \forall x \prec \overline{\alpha}\, A(x) \mid A(v) \in \mathsf{Sent}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}} \text{ and } \alpha < \varepsilon_\alpha\}$$

Theorem

$\mathrm{KFL}_{\varepsilon_\alpha} \subseteq \mathrm{R}_\alpha(\mathrm{KFL})$, *i.e.* $\mathrm{KFL}_{\varepsilon_\alpha}$ *is a subtheory of* α*-many iterations of Reflection over* KFL.

Lemma

$\mathrm{KFL}_{\varepsilon_{\alpha+1}} \subseteq \mathrm{R}(\mathrm{KFL}_{\varepsilon_\alpha})$

- As mentioned above, arithmetic in HYPE can deal with the Gentzen Jump formula. For $A(v) \in \mathcal{L}_{\mathrm{Tr}}^{\rightarrow}$,

$$\mathrm{PA}^{\rightarrow} \vdash \mathrm{Prog}(A) \rightarrow \mathrm{Prog}(A') \tag{1}$$

$$\mathrm{PA}^{\rightarrow} \vdash \forall\alpha(\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A', \alpha) \rightarrow \mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A, \omega^{\alpha}) \tag{2}$$

- One can generalize (2) to arbitrarily large (finite) iterations of the jump A'_k, so that we can prove:

$$\mathrm{PA}^{\rightarrow} \vdash \forall\alpha(\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A'_n, \alpha) \rightarrow \mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A, \omega^{\alpha}_n),$$

with $\omega^{\alpha}_0 = \alpha$, and $\omega^{\alpha}_{n+1} = \omega^{\omega^{\alpha}_n}$.

- We want $\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(< \varepsilon_{\alpha+1})$. So we can fix an $A \in \mathcal{L}_{\mathrm{Tr}}^{\rightarrow}$ and a $\gamma < \omega_n^{\varepsilon_\alpha+1} < \varepsilon_{\alpha+1}$ for some n, given than

$$\varepsilon_{\alpha+1} = \sup\{\omega_m^{\varepsilon_\alpha+1} \mid m \in \omega\}.$$

- It will then suffice to obtain $\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A_n, \varepsilon_\alpha)$. However, ε_α itself is the supremum of a sequence $\{f(n) \mid n \in \omega\}$. We have:

$$\mathrm{PA}^{\rightarrow} \vdash \forall x \prec \varepsilon_\alpha \exists y \prec \omega \; x \prec f(y)$$

$$\mathrm{PA}^{\rightarrow} \vdash \forall y \, \mathrm{Prov}_{\mathrm{KFL}_{\varepsilon_\alpha}}(\ulcorner \mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A_n, f(\overline{y})) \urcorner)$$

- Therefore, an application of Reflection for $\mathrm{KFL}_{\varepsilon_\alpha}$ yields $\mathrm{TI}_{\mathcal{L}_{\mathrm{Tr}}^{\rightarrow}}(A_n, f(y))$, as required.

Corollary

$\mathrm{R}_\alpha(\mathrm{KFL})$ *proves* $\mathrm{TI}_{\mathcal{L}_{\mathbb{N}}}(< \varphi_{\varepsilon_\alpha} 0)$.

To obtain an analogue of Leigh's for KFL – i.e. a proof of $\mathrm{R}_\alpha(\mathrm{KFL})$ in $\mathrm{KFL}_{\varepsilon_\alpha}$ – one would need to embed $\mathrm{KFL}_{\varepsilon_\alpha}$ in a suitable infinitary system, show (partial) *cut elimination* for this system, and then finally formalize the procedure within $\mathrm{KFL}_{\varepsilon_\alpha}$.

Unfortunately, HYPE seems to suffer from the same problems as the intermediate logic obtained by forcing constants domains in predicate intuitionistic logic. There are validities, e.g. the Fitting 'sequent'

$$\forall x(A \vee B(x)) \Rightarrow A, (\top \to \forall x B(x)),$$

which do not have a cut-free proof.

- Suppose one starts with justified belief in a reasonable mathematical theory T, and entertains an epistemic attitude towards Reflection and $\mathrm{UTB}^{\rightarrow}$ that is akin to the one entertained for *logical principles* (or anyway warranted in some non-deductive manner).
- Unlike the process developed in Horsten and Leigh (2015), Reflection can be understood in the more natural (and arguably less committing) form of *Global Reflection*.
- Already the interaction of the base theory, $\mathrm{UTB}^{\rightarrow}$, and reflection yields a theory that is as strong as the simple reflective closure of PA (KFL). From the truth-theoretic perspective, one is (explicitly) committed to strong principles of truth.
- The iteration of this process yields more and more instances of transfinite induction. The mathematical content of the process reaches the limits of predicativity.
- Unlike the process outlined in Fischer, Horsten, N. (2021), *both mathematical and truth-theoretic principles* are obtained in the form of *theorems*, and therefore are candidates for justified belief.