Modalizing full disquotation

Carlo Nicolai

Motivation and plan

Halbach (2001) and Horsten & Leigh (2016): recovering compositionality from *typed or positive disquotation* via reflection.

Motivation and plan

Halbach (2001) and Horsten & Leigh (2016): recovering compositionality from *typed or positive disquotation* via reflection.

Fischer, Nicolai, and Horsten (2017): recovering compositionality from *full disquotation*.

Motivation and plan

Halbach (2001) and Horsten & Leigh (2016): recovering compositionality from *typed or positive disquotation* via reflection.

Fischer, Nicolai, and Horsten (2017): recovering compositionality from *full disquotation*.

Both these projects are based on the following presuppostions:

1. disquotational principles have a superior semantic/modal status;
2. unraveling one's *implicit commitment on truth principles* via *reflection* is a sound strategy;

Motivation and plan

Halbach (2001) and Horsten & Leigh (2016): recovering compositionality from *typed or positive disquotation* via reflection.

Fischer, Nicolai, and Horsten (2017): recovering compositionality from *full disquotation*.

Both these projects are based on the following presuppostions:

1. disquotational principles have a superior semantic/modal status;
2. unraveling one's *implicit commitment on truth principles* via *reflection* is a sound strategy;

There seem to be reasons, if not to doubt 2, at least not to take it for granted. How can these projects look like starting only with 1?

Motivation and plan

Halbach (2001) and Horsten & Leigh (2016): recovering compositionality from *typed or positive disquotation* via reflection.

Fischer, Nicolai, and Horsten (2017): recovering compositionality from *full disquotation*.

Both these projects are based on the following presuppostions:

1. disquotational principles have a superior semantic/modal status;
2. unraveling one's *implicit commitment on truth principles* via *reflection* is a sound strategy;

There seem to be reasons, if not to doubt 2, at least not to take it for granted. How can these projects look like starting only with 1?

Slides are available at www.carlonicolai.org.

Let $\mathcal{L} = \{0, \mathrm{S}, +, \times, \exp\}$ and $\mathcal{L}_{\mathrm{Tr}} = \mathcal{L} \cup \{\mathrm{Tr}\}$.

$$\mathrm{TB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma \mid \sigma : \mathrm{Sent}_{\mathcal{L}}\}$$

$$\mathrm{PTB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma \mid \sigma : \mathrm{Sent}^{\mathrm{pos}}_{\mathcal{L}_{\mathrm{Tr}}}\}$$

$$\mathrm{RFN}(T) := T + \{\forall x(\mathrm{Pr}_T(\ulcorner\varphi(\dot{x})\urcorner) \to \varphi(x)) \mid \varphi(v) : \mathrm{Fml}_{\mathcal{L}_T}, T \supseteq \mathrm{EA}\}$$

Let $\mathcal{L} = \{0, \mathrm{S}, +, \times, \exp\}$ and $\mathcal{L}_{\mathrm{Tr}} = \mathcal{L} \cup \{\mathrm{Tr}\}$.

$$\mathrm{TB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma \mid \sigma : \mathrm{Sent}_{\mathcal{L}}\}$$

$$\mathrm{PTB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma \mid \sigma : \mathrm{Sent}^{\mathrm{pos}}_{\mathcal{L}_{\mathrm{Tr}}}\}$$

$$\mathrm{RFN}(T) := T + \{\forall x(\mathrm{Pr}_T(\ulcorner\varphi(\dot{x})\urcorner) \to \varphi(x)) \mid \varphi(v) : \mathrm{Fml}_{\mathcal{L}_T}, T \supseteq \mathrm{EA}\}$$

Halbach 2001, Cieśliński 2010

- $\mathrm{RFN}(\mathrm{TB}_0[\mathrm{EA}]) \vdash \mathrm{UTB}[\mathrm{PA}] (:= \mathrm{UTB})$
- $\mathrm{RFN}^2(\mathrm{TB}_0[\mathrm{EA}]) \vdash \mathrm{CT}[\mathrm{PA}] (:= \mathrm{CT})$
- $\mathrm{RFN}(\mathrm{PTB}_0[\mathrm{EA}])$ interprets KF.

For the negation-free language $\mathcal{L}_{\mathrm{Tr,F}}$ and σ^{D} the dual of σ:

$$\mathrm{TFB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma, \mathrm{F}\ulcorner\sigma^{D}\urcorner \leftrightarrow \sigma \mid \sigma \in \mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr,F}}}\}$$

For the negation-free language $\mathcal{L}_{\mathrm{Tr,F}}$ and σ^D the dual of σ:

$$\mathrm{TFB}_0[\mathrm{EA}] := \mathrm{EA} + \{\mathrm{Tr}\ulcorner\sigma\urcorner \leftrightarrow \sigma, \mathrm{F}\ulcorner\sigma^D\urcorner \leftrightarrow \sigma \mid \sigma \in \mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr,F}}}\}$$

Horsten & Leigh 2016

- $\mathrm{RFN}(\mathrm{TFB}_0[\mathrm{EA}]) \vdash \mathrm{UTFB}$
- $\mathrm{RFN}^2(\mathrm{TFB}_0[\mathrm{EA}]) \vdash \mathrm{KF}$

From now on I'll omit reference to EA and take this for granted in the context of classical theories.

Fact

The truth predicate of any $T \supseteq \mathrm{EA}$ in $\mathcal{L}_{\mathrm{Tr}}$ satisfying, for any $\sigma \in \mathcal{L}_{\mathrm{Tr}}$:

(1) $(\mathrm{Tr}\,\ulcorner \mathrm{Tr}\,\dot{x} \urcorner \leftrightarrow \mathrm{Tr}\,x) \wedge (\mathrm{Tr}\,\ulcorner \neg \mathrm{Tr}\,\dot{x} \urcorner \leftrightarrow \mathrm{Tr}\,\neg x)$

(2) $\mathrm{Pr}_{\mathrm{EA}_{\mathcal{L}_{\mathrm{Tr}}}}(\ulcorner \sigma \urcorner) \rightarrow \mathrm{Tr}\,\ulcorner \sigma \urcorner$

(3) for $\circ \in \{\wedge, \vee\}$, $\tau \in \mathcal{L}_{\mathrm{Tr}}$, $\mathrm{Tr}\,(\ulcorner \sigma \circ \tau \urcorner) \leftrightarrow (\mathrm{Tr}\,\ulcorner \sigma \urcorner \circ \mathrm{Tr}\,\ulcorner \tau \urcorner)$

is inconsistent.

Fact

The truth predicate of any $T \supseteq \mathrm{EA}$ in $\mathcal{L}_{\mathrm{Tr}}$ satisfying, for any $\sigma \in \mathcal{L}_{\mathrm{Tr}}$:

(1) $(\mathrm{Tr}\ulcorner \mathrm{Tr}\,\dot{x}\urcorner \leftrightarrow \mathrm{Tr}\,x) \wedge (\mathrm{Tr}\ulcorner \neg\mathrm{Tr}\,\dot{x}\urcorner \leftrightarrow \mathrm{Tr}\,\neg x)$

(2) $\mathrm{Pr}_{\mathrm{EA}_{\mathcal{L}_{\mathrm{Tr}}}}(\ulcorner\sigma\urcorner) \to \mathrm{Tr}\ulcorner\sigma\urcorner$

(3) for $\circ \in \{\wedge, \vee\}$, $\tau \in \mathcal{L}_{\mathrm{Tr}}$, $\mathrm{Tr}(\ulcorner\sigma \circ \tau\urcorner) \leftrightarrow (\mathrm{Tr}\ulcorner\sigma\urcorner \circ \mathrm{Tr}\ulcorner\tau\urcorner)$

is inconsistent.

The strategy of reflecting over a theory satisfying the above principles needs to cut the links between the standard of truth assumed in uniform reflection and the object linguistic notion of truth. There seem to be least two possible conclusions:

- Keep the metatheoretic notion of truth implicit in reflection and question the object linguistic one
- Keep the object linguistic notion of truth and question reflection

Following a strategy initiated by Halbach (2002) I 'rewrite' the results via a suitable modal predicate $\Box$ **closed under logical provability**:

Modalized disquotation: general pattern

Let $\mathcal{L}_P^{\Box} := \mathcal{L}_P \cup \{\Box\}$. For T a base truth system of the one considered, $\mathrm{M}^{\mathrm{k}}[T]$ is obtained by

- adding to it the claim 'all relevant Tr-biconditionals are boxed' – similarly for falsity biconditionals.
- if σ is an $\mathcal{L}_{\vec{P}}$-sentence provable in $\mathrm{M}^{\mathrm{k}}[T]$, then $\mathrm{M}^{\mathrm{k}}[T] \vdash \Box\ulcorner\sigma\urcorner$
- for $\varphi(v)$ a formula of $\mathcal{L}_{\vec{P}}$, $\mathrm{M}^{\mathrm{k}}[T]$ proves all instances of $\forall x(\Box\ulcorner\varphi(\dot{x})\urcorner \to \varphi(x))$

Example: $\mathrm{M^k[TB_0]}$

Extends TB_0 with:

- $(\forall x : \mathrm{Sent}_{\mathcal{L}^{\Box}_{\mathrm{Tr}}})(\mathrm{Pr}_{\mathrm{Cl}}(x) \to \Box x) \wedge$
  $(\forall x, y : \mathrm{Sent}_{\mathcal{L}^{\Box}_{\mathrm{Tr}}})(\Box(x \underset{\cdot}{\to} y) \wedge \Box x \to \Box y)$
- $(\forall \sigma : \mathcal{L})(\Box \ulcorner \mathrm{Tr} \ulcorner \dot{\sigma} \urcorner \leftrightarrow \sigma \urcorner)$
  i.e. $(\forall x : \mathrm{Sent}_{\mathcal{L}})(\Box(\mathrm{sub}(\underset{\cdot}{\leftrightarrow}(\underset{\cdot}{\mathrm{Tr}}\, v, x), v, \mathrm{num}\, x)))$
- for σ a sentence of $\mathcal{L}_{\mathrm{Tr}}$, if $\mathrm{M^k[TB_0]} \vdash \sigma$, then $\mathrm{M^k[TB_0]} \vdash \Box \ulcorner \sigma \urcorner$

Proposition (Typed biconditionals)

- $\mathrm{M}^{\mathrm{k}}[\mathrm{TB}_0] \vdash \mathrm{Ind}(\mathcal{L}^{\square}_{\mathrm{Tr}})$
- $\mathrm{M}^{\mathrm{k}}[\mathrm{TB}_0]$ proves all Tarski biconditionals for $\mathcal{L}$-formulas with free variables (UTB) and the (typed) compositional clauses for the logical connectives – e.g. $\forall x : \mathsf{Sent}_{\mathcal{L}}(\mathrm{Tr}\,\underset{\cdot}{\neg} x \leftrightarrow \neg \mathrm{Tr}\, x)$;
- $\mathrm{M}^{\mathrm{k}}[\mathrm{UTB}_0]$ proves the compostional clauses for quantifiers – e.g. $\forall v(\forall x : \mathrm{Fml}^1_{\mathcal{L}})(\mathrm{Tr}\,\underset{\cdot}{\forall} v x \leftrightarrow \forall y\, \mathrm{Tr}\, x(\dot{y}/v))$; that is, it includes CT. Futhermore, it's reducible to ACA.

Proposition (Typed biconditionals)

- $\mathrm{M^k[TB_0]} \vdash \mathrm{Ind}(\mathcal{L}^{\Box}_{\mathrm{Tr}})$
- $\mathrm{M^k[TB_0]}$ proves all Tarski biconditionals for $\mathcal{L}$-formulas with free variables (UTB) and the (typed) compositional clauses for the logical connectives – e.g. $\forall x : \mathsf{Sent}_{\mathcal{L}}(\mathrm{Tr}\,\underset{\cdot}{\neg} x \leftrightarrow \neg \mathrm{Tr}\, x)$;
- $\mathrm{M^k[UTB_0]}$ proves the compostional clauses for quantifiers – e.g. $\forall v(\forall x : \mathrm{Fml}^1_{\mathcal{L}})(\mathrm{Tr}\,\underset{\cdot}{\forall} v x \leftrightarrow \forall y\, \mathrm{Tr}\, x(\dot{y}/v))$; that is, it includes CT. Futhermore, it's reducible to ACA.

Proposition (Type-free biconditionals)

- $\mathrm{M^k[TFB_0]}$ includes UTFB; therefore it includes PUTB and there is an $\mathcal{L}$-embedding of KF in it;
- $\mathrm{M^k[UTFB]}$ includes KF and is reducible to it.

Necessity does not imply truth

Any truth predicate of any *modalized* extension of EA in $\mathcal{L}^{\Box}_{\mathrm{Tr}}$ satisfying, for $\sigma, \tau \in \mathcal{L}_{\mathrm{Tr}}$,

$$\mathrm{Tr}\,(\ulcorner \sigma \circ \tau \urcorner) \leftrightarrow \mathrm{Tr}\,\ulcorner \sigma \urcorner \circ \mathrm{Tr}\,\ulcorner \tau \urcorner \text{ with } \circ \in \{\wedge, \vee\}$$

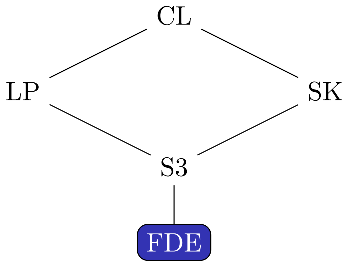
$$(\mathrm{Tr}\,\ulcorner \mathrm{Tr}\,\dot{x} \urcorner \leftrightarrow \mathrm{Tr}\,x) \wedge (\mathrm{Tr}\,\ulcorner \neg \mathrm{Tr}\,\dot{x} \urcorner \leftrightarrow \mathrm{Tr}\,\neg x)$$

$$\Box \ulcorner \sigma \urcorner \to \mathrm{Tr}\,\ulcorner \sigma \urcorner$$

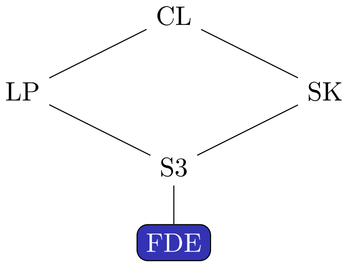
is inconsistent.

There is also a fundamental philosophical worry concerning the notion of analyticity at stake. I assume some work must be done there: analyticity does not always entail necessity (it would depend on the story we tell concerning the justification of the biconditionals).

Modalizing Full Disquotation: Logic and Syntax



Modalizing Full Disquotation: Logic and Syntax



The theory Basic

Working in $\mathcal{L}_{\mathrm{Tr}}$, we extend a (two-sided) sequent calculus formulation of FDE with the basic axioms of $\mathrm{I}\Delta_0 + \Omega_1$ and a $\Delta_0(\omega_1)$ induction rule for $\mathcal{L}$:

$$\frac{\Gamma, \varphi(x) \Rightarrow \varphi(\mathrm{S}x), \Delta}{\Gamma, \varphi(0) \Rightarrow \varphi(y), \Delta}$$

for y not free in $\varphi(0), \Gamma, \Delta$. The restriction of Basic to $\mathcal{L}$ is simply $\mathrm{I}\Delta_0 + \Omega_1$.

Modalizing Full Disquotation: Truth and Necessity

In setting up the system, a modality applying to single sentences or a modality applying to single sequents won't be enough:

Modalizing full disquotation: General Template

Extend Basic in $\mathcal{L}^{\boxdot}_{\mathrm{Tr}}$ with:

- $\boxdot(x, y)$ includes logically admissible rules and it's closed under FDE-derivability;
- the relevant disquotational sequents are in the extension of $\boxdot$;
- $$\frac{\Rightarrow \boxdot(\ulcorner\Gamma(\dot{x})\Rightarrow\Delta(\dot{x})\urcorner / \ulcorner\Pi(\dot{x})\Rightarrow\Sigma(\dot{x})\urcorner) \qquad \Gamma(x) \Rightarrow \Delta(x)}{\Pi(x) \Rightarrow \Sigma(x)}$$
- if $\Gamma \Rightarrow \Delta \in \mathcal{L}^{\mathrm{Tr}}$ is derivable, then $\Rightarrow \boxdot\ulcorner\Gamma \Rightarrow \Delta\urcorner$ is — where $\boxdot(x) :\leftrightarrow \boxdot(\underset{.}{\Rightarrow}\top, x)$.

Substitution and numerals are now dealt with ω_1-style.

Truth and Necessity again. . .

The move to the sub-classical setting enables one to restore the link between necessity and truth. Let's consider the following extensions of Basic with the principles

$$(\mathrm{BT}_0) \qquad \mathrm{Tr}\ulcorner\sigma\urcorner \Rightarrow \sigma, \qquad \sigma \Rightarrow \mathrm{Tr}\ulcorner\sigma\urcorner;$$

$$(\mathrm{UBT}_0) \qquad \mathrm{Tr}\ulcorner\varphi(\dot{x})\urcorner \Rightarrow \varphi(x), \qquad \varphi(x) \Rightarrow \mathrm{Tr}\ulcorner\varphi(\dot{x})\urcorner.$$

for $\sigma, \varphi(v) \in \mathcal{L}_{\mathrm{Tr}}$.

Truth and Necessity again. . .

The move to the sub-classical setting enables one to restore the link between necessity and truth. Let's consider the following extensions of Basic with the principles

$$(\mathrm{BT}_0) \qquad \mathrm{Tr}\ulcorner\sigma\urcorner \Rightarrow \sigma, \qquad \sigma \Rightarrow \mathrm{Tr}\ulcorner\sigma\urcorner;$$

$$(\mathrm{UBT}_0) \qquad \mathrm{Tr}\ulcorner\varphi(\dot{x})\urcorner \Rightarrow \varphi(x), \qquad \varphi(x) \Rightarrow \mathrm{Tr}\ulcorner\varphi(\dot{x})\urcorner.$$

for $\sigma, \varphi(v) \in \mathcal{L}_{\mathrm{Tr}}$.

Lemma

Over any $T \supseteq \mathrm{UBT}_0$ in which at least all instances of BT_0 are boxed, the following rules are equivalent:

$$\frac{\Rightarrow \mathrm{Sent}_{\mathcal{L}^{\boxdot}_{\mathrm{Tr}}}(x), \boxdot(\dot{\Rightarrow} x)}{\Rightarrow T(x)} \qquad \frac{\Rightarrow \boxdot(\ulcorner \Rightarrow \varphi(\dot{x})\urcorner)}{\Rightarrow \varphi(x)}$$

Therefore we do not need to worry about the previous incoherence.

Modalization of full disquotation: an overview

Proposition

- $\mathrm{M}[\mathrm{BT}_0]$ proves
  - all instances of UBT_0
  - the full induction rule for $\mathcal{L}_{\mathrm{Tr}}^{\boxdot}$ (and transfinite induction for $\mathcal{L}_{\mathrm{Tr}}^{\boxdot}$ up to and not including ω^ω)
  - propositional compositional sequents such as

$$\mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}}(x), \mathrm{Tr}\,\dot{\neg}x \Rightarrow \neg\mathrm{Tr}\,x$$

Modalization of full disquotation: an overview

Proposition

- $\mathrm{M}[\mathrm{BT}_0]$ proves
  - all instances of UBT_0
  - the full induction rule for $\mathcal{L}^{\boxdot}_{\mathrm{Tr}}$ (and transfinite induction for $\mathcal{L}^{\boxdot}_{\mathrm{Tr}}$ up to and not including ω^{ω})
  - propositional compositional sequents such as

$$\mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}}(x), \mathrm{Tr}\,\dot{\neg} x \Rightarrow \neg \mathrm{Tr}\, x$$

- $\mathrm{M}[\mathrm{UBT}]$ proves
  - Compositional sequents for quantifiers, e.g.

$$\mathrm{Sent}_{\mathcal{L}_{\mathrm{Tr}}}(x), \mathrm{Tr}\,\dot{\forall} v x \Rightarrow \forall y\, \mathrm{Tr}\, x(\dot{y}/v)$$

  - Transfinite induction for the language with the truth predicate up to (including ω^{ω}) – arithmetical induction up to $\phi_{\omega}0$ – more than the version of Kripke-Feferman in FDE, PKF from Halbach and Horsten 2006.

Summing up

If one does not accept reflection without an argument, one can employ the alleged modal status of disquotational sentences.

Summing up

If one does not accept reflection without an argument, one can employ the alleged modal status of disquotational sentences.

Modalized disquotational theories over classical logic face two main drawbacks: the arbitrariness of the choice of the biconditionals and an ill-behaved modality.

Summing up

If one does not accept reflection without an argument, one can employ the alleged modal status of disquotational sentences.

Modalized disquotational theories over classical logic face two main drawbacks: the arbitrariness of the choice of the biconditionals and an ill-behaved modality.

Modalizing full disquotation does not suffer from the problems above: their semantic and modal notions are coherent but the resulting theories lack a satisfactory conditional and, perhaps more importantly, *is* $\boxdot$ *a modal predicate*?