The Phonological Conundrum in Formal Theories of Truth and Modalities

Carlo Nicolai



Jerusalem, June 5, 2019 (slides at carlonicolai.github.io)

\sim	\sim	01	100	 m?

Every integer greater than 2 is equal to the sum of three primes $(\forall n > 2)(\exists m, k, l : Prime)(n = m + k + l)$

$$(\forall n > 2)(\exists m, k, l : Prime)(n = m + k + l)$$

All axioms of Peano Arithmetic are true

$$(\forall n > 2)(\exists m, k, l : Prime)(n = m + k + l)$$

All axioms of Peano Arithmetic are true

$$(\forall \varphi \in \mathcal{L}_{\mathbb{N}})(\operatorname{Ax}_{\operatorname{PA}}(\varphi) \to \operatorname{Tr} \varphi)$$

$$(\forall n > 2)(\exists m, k, l : Prime)(n = m + k + l)$$

All axioms of Peano Arithmetic are true

$$(\forall \varphi \in \mathcal{L}_{\mathbb{N}})(\operatorname{Ax}_{\operatorname{PA}}(\varphi) \to \operatorname{Tr} \varphi)$$

Any arithmetical property that is true of 0, and that is true of n+1 if it is true of n, is true of any natural number.

$$(\forall n > 2)(\exists m, k, l : Prime)(n = m + k + l)$$

All axioms of Peano Arithmetic are true

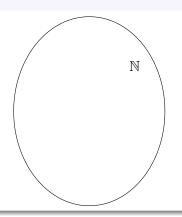
$$(\forall \varphi \in \mathcal{L}_{\mathbb{N}})(\operatorname{Ax}_{\operatorname{PA}}(\varphi) \to \operatorname{Tr} \varphi)$$

Any arithmetical property that is true of 0, and that is true of n+1 if it is true of n, is true of any natural number.

$$\operatorname{Tr} \lceil \varphi(0) \rceil \wedge \forall x (\operatorname{Tr} \lceil \varphi(\dot{x}) \rceil \to \operatorname{Tr} \lceil \varphi(\dot{x}+1) \rceil) \to \forall x \operatorname{Tr} \lceil \varphi(\dot{x}) \rceil$$

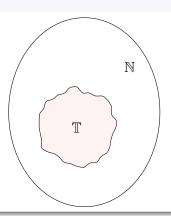
The standard setting

 $\blacktriangleright \mathcal{L}_{Tr} := \mathcal{L}_{\mathbb{N}} \cup \{Tr\}$



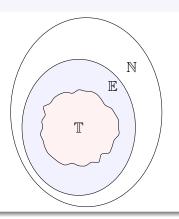
The standard setting

 $\blacktriangleright \mathcal{L}_{Tr} := \mathcal{L}_{\mathbb{N}} \cup \{Tr\}$



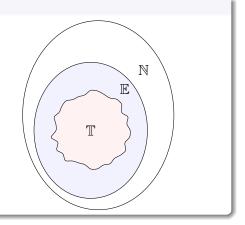
The standard setting

 $\blacktriangleright \mathcal{L}_{Tr} := \mathcal{L}_{\mathbb{N}} \cup \{Tr\}$



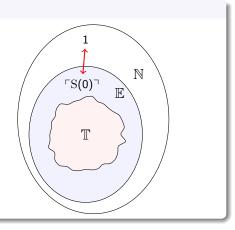
The standard setting

- One sort of quantification
- One kind of induction



The standard setting

- $\blacktriangleright \mathcal{L}_{Tr} := \mathcal{L}_{\mathbb{N}} \cup \{Tr\}$
- ► One sort of quantification
- One kind of induction
- Some implicit bridge principles



some implications/1

In the standard setting, we are not in general a theory of truth for languages, but only for a restricted range of theories.

In the standard setting, we are not in general a theory of truth for languages, but only for a restricted range of theories.

Example

The language of mereology $\{\circ\}$. Virtually all reasonable, pure mereological theories are decidable. Therefore, there is no hope of formulating a reasonable syntax theory in them.

In the standard setting, we are not in general a theory of truth for languages, but only for a restricted range of theories.

Example

The language of mereology $\{\circ\}$. Virtually all reasonable, pure mereological theories are decidable. Therefore, there is no hope of formulating a reasonable syntax theory in them.

In a more general setting, what we learn about truth is also more general (cf. Tarski 1936).

some implications/2

Philosophical analysis often requires to be able to distinguish between different patterns of reasoning.

Philosophical analysis often requires to be able to distinguish between different patterns of reasoning.

Is truth explanatory?

Shapiro 1998, Cieśliński 2017:

▶ Yes, if the theory of truth is non-conservative,

Philosophical analysis often requires to be able to distinguish between different patterns of reasoning.

Is truth explanatory?

Shapiro 1998, Cieśliński 2017:

- Yes, if the theory of truth is non-conservative,
- No, if it is conservative.

Philosophical analysis often requires to be able to distinguish between different patterns of reasoning.

Is truth explanatory?

Shapiro 1998, Cieśliński 2017:

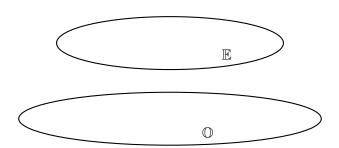
- Yes, if the theory of truth is non-conservative,
- No, if it is conservative.

Analyses of this sort rely essentially on the way in which induction is presented in the standard setting.

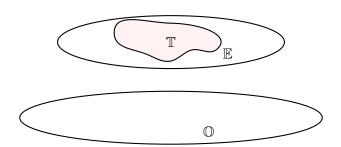




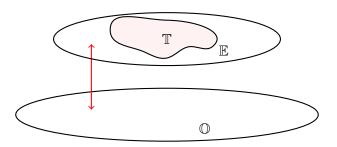
disentangled syntax

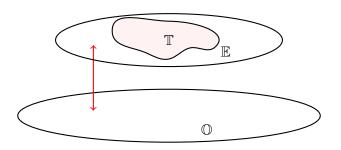


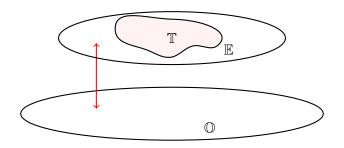
disentangled syntax



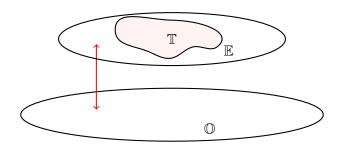
disentangled syntax





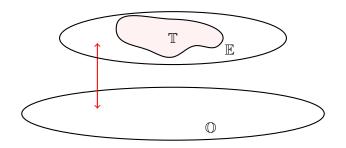


Axioms of the subject matter O



Axioms of the subject matter O

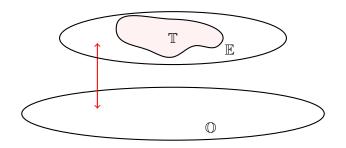
Axioms for the syntax theory S



Axioms of the subject matter O

Axioms for the syntax theory S

Truth, or better satisfaction axioms



Axioms of the subject matter O

Axioms for the syntax theory S

Truth, or better satisfaction axioms

Axioms for the bridge principles

Such desiderata are met, for instance, by Buss's theory of p-time computability called S_2^1 (or Ferreira's string theoretic variant Σ_1^b -NIA):

Such desiderata are met, for instance, by Buss's theory of p-time computability called S_2^1 (or Ferreira's string theoretic variant Σ_1^b -NIA):

▶ the smash function $x, y \mapsto 2^{|x| \cdot |y|}$

Such desiderata are met, for instance, by Buss's theory of p-time computability called S_2^1 (or Ferreira's string theoretic variant Σ_1^b -NIA):

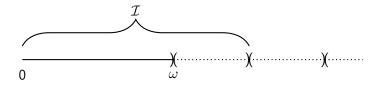
- ▶ the smash function $x, y \mapsto 2^{|x|\cdot|y|}$
- ▶ the Σ_1^1 -PIND induction schema:

$$\varphi(0) \land \forall x (\varphi(\lfloor \frac{1}{2}x \rfloor) \to \varphi(x)) \to \forall x \varphi(x) \text{ for } \varphi \in \Sigma_1^b$$

► Smooth formalisation of syntactic notions up to Gödel's II

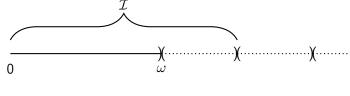
- Smooth formalisation of syntactic notions up to Gödel's II
- ► Close to the theoretical lower bound: Robinson's Q interprets S_2^1 on a cut.

- Smooth formalisation of syntactic notions up to Gödel's II
- ▶ Close to the theoretical lower bound: Robinson's Q interprets S_2^1 on a cut.



$$\forall x, y \in \mathcal{I} \ \exists z \in \mathcal{I} : z = x \# y$$

- Smooth formalisation of syntactic notions up to Gödel's II
- ▶ Close to the theoretical lower bound: Robinson's Q interprets S_2^1 on a cut.



$$\forall x, y \in \mathcal{I} \ \exists z \in \mathcal{I} : z = x \# y$$

lt is finitely axiomatisable.

$$\forall \alpha \forall o \forall n \exists \alpha \forall m ((m \neq n \rightarrow \alpha(n) = \beta(n)) \land \beta(m) = o)$$
 (SEQ)

$$\forall \alpha \forall o \forall n \exists \alpha \forall m ((m \neq n \rightarrow \alpha(n) = \beta(n)) \land \beta(m) = o)$$
 (SEQ)

This guarantees that the we have at our disposal v_m -variants when needed.

$$\forall \alpha \forall o \forall n \exists \alpha \forall m \big((m \neq n \rightarrow \alpha(n) = \beta(n)) \land \beta(m) = o \big) \quad \text{(SEQ)}$$

This guarantees that the we have at our disposal v_m -variants when needed.

If constants and/or function symbols are present, we need a denotation function:

$$\forall \alpha \forall o \forall n \exists \alpha \forall m ((m \neq n \rightarrow \alpha(n) = \beta(n)) \land \beta(m) = o)$$
 (SEQ)

This guarantees that the we have at our disposal v_m -variants when needed.

If constants and/or function symbols are present, we need a denotation function:

$$\alpha, \lceil c \rceil \mapsto c$$

$$\forall \alpha \forall o \forall n \exists \alpha \forall m ((m \neq n \rightarrow \alpha(n) = \beta(n)) \land \beta(m) = o)$$
 (SEQ)

This guarantees that the we have at our disposal v_m -variants when needed.

If constants and/or function symbols are present, we need a denotation function:

$$\alpha, \lceil c \rceil \mapsto c$$

 $\alpha, \lceil f(v_n) \rceil \mapsto f(\alpha(n))$

what we know so far

$\mathrm{ut}[O]$

▶ Axioms for a RE, first-order object theory O in \mathcal{L}_O ;

$\mathrm{ut}[O]$

- ightharpoonup Axioms for a RE, first-order object theory O in \mathcal{L}_O ;
- ► $\operatorname{Sat}(\alpha, \lceil \varphi(v_i) \rceil) \leftrightarrow \varphi(\alpha(i))$ for $\varphi(v_i) \in \mathcal{L}_O$

ut[0]

- ▶ Axioms for a RE, first-order object theory O in \mathcal{L}_O ;
- ► $\operatorname{Sat}(\alpha, \lceil \varphi(v_i) \rceil) \leftrightarrow \varphi(\alpha(i))$ for $\varphi(v_i) \in \mathcal{L}_O$
- ▶ S_2^1 with Σ_1^b -PIND induction extended to Sat

ut[*O*]

- ▶ Axioms for a RE, first-order object theory O in \mathcal{L}_O ;
- ► $\operatorname{Sat}(\alpha, \lceil \varphi(v_i) \rceil) \leftrightarrow \varphi(\alpha(i))$ for $\varphi(v_i) \in \mathcal{L}_O$
- ▶ S_2^1 with Σ_1^b -PIND induction extended to Sat
- ► (SEQ)

ut[*O*]

- ▶ Axioms for a RE, first-order object theory O in \mathcal{L}_O ;
- ► $\operatorname{Sat}(\alpha, \lceil \varphi(v_i) \rceil) \leftrightarrow \varphi(\alpha(i))$ for $\varphi(v_i) \in \mathcal{L}_O$
- ▶ S_2^1 with Σ_1^b -PIND induction extended to Sat
- ► (SEQ)

Lemma

If O has a good notion of sequence, $\operatorname{ut}[O]$ is locally interpretable (and therefore conservative) over O.

ut[0]

- Axioms for a RE, first-order object theory O in \mathcal{L}_O ;
- ► $\operatorname{Sat}(\alpha, \lceil \varphi(v_i) \rceil) \leftrightarrow \varphi(\alpha(i))$ for $\varphi(v_i) \in \mathcal{L}_O$
- ▶ S_2^1 with Σ_1^b -PIND induction extended to Sat
- ► (SEQ)

Lemma

If O has a good notion of sequence, $\operatorname{ut}[O]$ is locally interpretable (and therefore conservative) over O.

Proof. For the k-many relevant formulas appearing in instances of disquotation, define:

$$\operatorname{Sat}(x,y) : \leftrightarrow \bigvee_{n=1}^{k} \left(y = \overline{\varphi_n(v_i)} \wedge \varphi_n(x(i)) \right)$$

$ct[O]^-$

One restricts Σ_1^b -PIND to $\mathcal{L}_{\mathbb{N}}$, and adds the axioms:

$$Sat(\alpha, \lceil R(v_1, ..., v_n) \rceil) \leftrightarrow R(\alpha(1), ..., \alpha(n)) \text{ for } R \in \mathcal{L}_O$$

$$Sat(\alpha, \lceil \neg \varphi \rceil) \leftrightarrow \neg Sat(\alpha, \lceil \varphi \rceil)$$

$$Sat(\alpha, \lceil \varphi \land \psi \rceil) \leftrightarrow Sat(\alpha, \lceil \varphi \rceil) \land Sat(\alpha, \lceil \psi \rceil)$$

$$Sat(\alpha, \lceil \forall v_i \varphi \rceil) \leftrightarrow \forall \beta \ (\forall j (j \neq i \rightarrow \alpha(j) = \beta(j)) \rightarrow Sat(\beta, \lceil \varphi \rceil))$$

ct[O]⁻

One restricts Σ_1^b -PIND to $\mathcal{L}_{\mathbb{N}}$, and adds the axioms:

$$Sat(\alpha, \lceil R(v_1, ..., v_n) \rceil) \leftrightarrow R(\alpha(1), ..., \alpha(n)) \text{ for } R \in \mathcal{L}_O$$

$$Sat(\alpha, \lceil \neg \varphi \rceil) \leftrightarrow \neg Sat(\alpha, \lceil \varphi \rceil)$$

$$Sat(\alpha, \lceil \varphi \wedge \psi \rceil) \leftrightarrow Sat(\alpha, \lceil \varphi \rceil) \wedge Sat(\alpha, \lceil \psi \rceil)$$

$$Sat(\alpha, \lceil \forall v_i \varphi \rceil) \leftrightarrow \forall \beta \ (\forall j (j \neq i \rightarrow \alpha(j) = \beta(j)) \rightarrow Sat(\beta, \lceil \varphi \rceil))$$

ct[O]

From $ct[0]^-$, Σ_1^b -PIND is extended to the entire language.

ct[O]-

One restricts Σ_1^b -PIND to $\mathcal{L}_{\mathbb{N}}$, and adds the axioms:

$$Sat(\alpha, \lceil R(v_1, ..., v_n) \rceil) \leftrightarrow R(\alpha(1), ..., \alpha(n)) \text{ for } R \in \mathcal{L}_O$$

$$Sat(\alpha, \lceil \neg \varphi \rceil) \leftrightarrow \neg Sat(\alpha, \lceil \varphi \rceil)$$

$$Sat(\alpha, \lceil \varphi \wedge \psi \rceil) \leftrightarrow Sat(\alpha, \lceil \varphi \rceil) \wedge Sat(\alpha, \lceil \psi \rceil)$$

$$Sat(\alpha, \lceil \forall v_i \varphi \rceil) \leftrightarrow \forall \beta \ (\forall j (j \neq i \rightarrow \alpha(j) = \beta(j)) \rightarrow Sat(\beta, \lceil \varphi \rceil))$$

ct[O]

From ct[O]⁻, Σ_1^b -PIND is extended to the entire language.

ct[O]+

The extended Σ_1^b -PIND is replaced with $\exists \Delta_1^b$ -induction

Lemma

▶ If O is finitely axiomatised, $ct[O]^+$ proves the Global Reflection Principle:

$$\forall \varphi \in \mathcal{L}_{\mathcal{O}}(\operatorname{Prov}_{\mathcal{O}}(\varphi) \to \forall \alpha \operatorname{Sat}(\alpha, \varphi))$$

Lemma

▶ If O is finitely axiomatised, $ct[O]^+$ proves the Global Reflection Principle:

$$\forall \varphi \in \mathcal{L}_{\mathcal{O}}(\operatorname{Prov}_{\mathcal{O}}(\varphi) \to \forall \alpha \operatorname{Sat}(\alpha, \varphi))$$

► For O infinitely (or schematically) axiomatised, one needs to add the claim:

$$\forall \varphi \in \mathcal{L}_O(Ax_O(\varphi) \to \forall \alpha \operatorname{Sat}(\alpha, \varphi))$$

Lemma

▶ If O is finitely axiomatised, ct[O]⁺ proves the Global Reflection Principle:

$$\forall \varphi \in \mathcal{L}_O(\text{Prov}_O(\varphi) \to \forall \alpha \operatorname{Sat}(\alpha, \varphi))$$

► For O infinitely (or schematically) axiomatised, one needs to add the claim:

$$\forall \varphi \in \mathcal{L}_O(\operatorname{Ax}_O(\varphi) \to \forall \alpha \operatorname{Sat}(\alpha, \varphi))$$

Corollary

With the conditions above, $ct[O]^+$ proves Con(O).

conservativeness

Proposition

 ${
m ct}[{\it O}], {
m ct}[{\it O}]^+$ are conservative extensions of ${\it O}.$

 $ct[O], ct[O]^+$ are conservative extensions of O.

Proof. Any model $\mathcal{M} \models O$ can be expanded to a model $\mathcal{M}' \models \operatorname{ct}[O]$:

 $ct[O], ct[O]^+$ are conservative extensions of O.

Proof. Any model $\mathcal{M} \models O$ can be expanded to a model $\mathcal{M}' \models \operatorname{ct}[O]$:

Syntax can be standardly interpreted

 $ct[O], ct[O]^+$ are conservative extensions of O.

Proof. Any model $\mathcal{M} \models O$ can be expanded to a model $\mathcal{M}' \models \operatorname{ct}[O]$:

- Syntax can be standardly interpreted
- ▶ Sequences are functions $f: \operatorname{Var}_{\mathcal{L}_Q}^{\mathbb{E}} \longrightarrow M$

 $ct[O], ct[O]^+$ are conservative extensions of O.

Proof. Any model $\mathcal{M} \models O$ can be expanded to a model $\mathcal{M}' \models \operatorname{ct}[O]$:

- Syntax can be standardly interpreted
- ▶ Sequences are functions $f: \operatorname{Var}_{\mathcal{L}_Q}^{\mathbb{E}} \longrightarrow M$

interpretability/1

Suppose *O* interprets Q:

Proposition

▶ If O is finitely axiomatised, $ct[O]^-$ is mutually interpretable with Q + Con(U)

Proposition

- ▶ If O is finitely axiomatised, $ct[O]^-$ is mutually interpretable with Q + Con(U)
- ▶ If O is schematically axiomatised, then $ct[O]^-+$ all axioms of O are true' is mutually interpretable with Q + Con(O).

Proposition

- ▶ If O is finitely axiomatised, $ct[O]^-$ is mutually interpretable with Q + Con(U)
- ▶ If O is schematically axiomatised, then $ct[O]^-+$ all axioms of O are true' is mutually interpretable with Q + Con(O).

Proof. I consider the first claim:

Proposition

- ▶ If O is finitely axiomatised, $ct[O]^-$ is mutually interpretable with Q + Con(U)
- ▶ If O is schematically axiomatised, then $ct[O]^-+$ all axioms of O are true' is mutually interpretable with Q + Con(O).

Proof. I consider the first claim:

▶ First, one proves the consistency of O on a $ct[O]^-$ -definable initial segment of the S_2^1 -numbers of O.

Proposition

- ▶ If O is finitely axiomatised, $ct[O]^-$ is mutually interpretable with Q + Con(U)
- ▶ If O is schematically axiomatised, then $ct[O]^-+$ all axioms of O are true' is mutually interpretable with Q + Con(O).

Proof. I consider the first claim:

- ▶ First, one proves the consistency of O on a $ct[O]^-$ -definable initial segment of the S_2^1 -numbers of O.
- ▶ Second, one employs a miniaturised version of the Henkin-construction: in Q + Con(O), one constructs a term model for O coming with a satisfaction predicate that satisfies $ct[O]^-$.

interpretability/2

Proposition

Let O be sequential, then $\mathrm{ut}[O]$ does not interpret $\mathrm{Q}+\mathrm{Con}(O)$.

interpretability/2

Proposition

Let O be sequential, then ut[O] does not interpret Q + Con(O).

Proof. If it did, O would interpret $S_2^1 + \operatorname{Con}(O)$, contradicting Pudlák's version of Gödel's second incompleteness theorem.

what does this say about truth?

▶ Philosophy is all about fine distinctions

what does this say about truth?

- Philosophy is all about fine distinctions
- ► Granted: an adequate theory of truth over a fairly arbitrary object theory *O* either proves or interprets a consistency statement for *O*.

- Philosophy is all about fine distinctions
- ▶ Granted: an adequate theory of truth over a fairly arbitrary object theory O either proves or interprets a consistency statement for O.
- However, in the present setting it is fairly clear that what is proved or interpreted is a syntactic claim that does not belong to the subject matter we are reasoning about.

- ▶ Philosophy is all about fine distinctions
- ▶ Granted: an adequate theory of truth over a fairly arbitrary object theory O either proves or interprets a consistency statement for O.
- ▶ However, in the present setting it is fairly clear that what is proved or interpreted is a syntactic claim that does not belong to the subject matter we are reasoning about.
- ► This seems compatible with the kinds of metatheoretic explanation that truth should provide.

doing without typing

lifting type restrictions

The setting just presented is essentially typed. Is there a type-free version of it?

Perhaps a Kripkean construction will do. For instance:

-
$$X_0 = \{(f, \varphi) \mid \varphi \in \mathcal{L}_O, f : \operatorname{Var}_{\mathcal{L}_O}^{\mathbb{E}} \to \mathbb{O}, \mathbb{O} \models_{k3}^f \varphi\}$$

Perhaps a Kripkean construction will do. For instance:

- $X_0 = \{(f, \varphi) \mid \varphi \in \mathcal{L}_O, f \colon \mathrm{Var}_{\mathcal{L}_O}^{\mathbb{E}} \to \mathbb{O}, \mathbb{O} \models_{k3}^f \varphi\}$
- $(\alpha, \operatorname{Sat}(v_0, v_1)) \in X_{\alpha+1}$ if $(\alpha(v_0), \alpha(v_1)) \in X_{\alpha}$

Perhaps a Kripkean construction will do. For instance:

- $-X_0=\{(f,\varphi)\mid \varphi\in\mathcal{L}_O, f\colon \mathrm{Var}_{\mathcal{L}_O}^{\mathbb{E}}\to\mathbb{O}, \mathbb{O}\vDash_{k3}^f\varphi\}$
- $(\alpha, \operatorname{Sat}(v_0, v_1)) \in X_{\alpha+1}$ if $(\alpha(v_0), \alpha(v_1)) \in X_{\alpha}$
- and so on for propositional (positive!) connectives.

Perhaps a Kripkean construction will do. For instance:

- $-\ X_0 = \{ (f,\varphi) \mid \varphi \in \mathcal{L}_O, f \colon \mathrm{Var}_{\mathcal{L}_O}^{\mathbb{E}} \to \mathbb{O}, \mathbb{O} \vDash_{k3}^f \varphi \}$
- $(\alpha, \operatorname{Sat}(v_0, v_1)) \in X_{\alpha+1}$ if $(\alpha(v_0), \alpha(v_1)) \in X_{\alpha}$
- and so on for propositional (positive!) connectives.

This toy construction reaches a fixed point at ω , but is is crucial that now sequences are mappings $\mathbb{E} \cup SEQ \to \mathbb{E} \cup SEQ \cup \mathbb{O}!$