Deflationism and Axiomatic Theories of Truth

Proof theoretic and model theoretic conservativities

Stella Moon

ILLC, UvA

2nd October 2015

Deflationism

- Truth is insubstantial so it does not carry any ontological weight.

Deflationism

- Truth is insubstantial so it does not carry any ontological weight.
- Deflationists are interested how truth works, rather than what it is.

Deflationism

- Truth is insubstantial so it does not carry any ontological weight.
- Deflationists are interested how truth works, rather than what it is.
- *It is true that snow is white iff snow is white.*

Deflationism

- Truth is insubstantial so it does not carry any ontological weight.
- Deflationists are interested how truth works, rather than what it is.
- *It is true that snow is white iff snow is white.*
- Motivated by Tarski's biconditionals: for any sentence ϕ

$$T(\phi) \leftrightarrow \phi.$$

Proof theoretic and model theoretic conservativities

- Deflationists desire our extended theory to be conservative over our base theory.

Proof theoretic and model theoretic conservativities

- Deflationists desire our extended theory to be conservative over our base theory.
- There are two notions of conservativities: for models and for theories.

Proof theoretic and model theoretic conservativities

Definition (Proof theoretic conservativity)

Let Γ *be a* $\mathcal{L}$*-theory and* Γ' *be a* $\mathcal{L}'$*-theory extending* Γ*, that is* $\mathcal{L}' \supseteq \mathcal{L}$*, such that* $\Gamma' \supseteq \Gamma$*.* Γ' *is proof theoretically conservative over* Γ *if for any* $\mathcal{L}$*-setenece* θ*,* $\Gamma' \vdash \theta$*, we have that* $\Gamma \vdash \theta$*.*

Proof theoretic and model theoretic conservativities

Definition (Proof theoretic conservativity)

Let Γ *be a* $\mathcal{L}$*-theory and* Γ' *be a* $\mathcal{L}'$*-theory extending* Γ*, that is* $\mathcal{L}' \supseteq \mathcal{L}$*, such that* $\Gamma' \supseteq \Gamma$*.* Γ' *is proof theoretically conservative over* Γ *if for any* $\mathcal{L}$*-setenece* θ, $\Gamma' \vdash \theta$*, we have that* $\Gamma \vdash \theta$*.*

Definition (Model theoretic conservativity)

Let Γ *be an* $\mathcal{L}$*-theory and* Γ' *be an* $\mathcal{L}'$*-theory extending* Γ*.* Γ' *is model-theoretically conservative over* Γ *if any model of* Γ *can be expanded to a model of* Γ'*.*

Peano Arithmetic

- Axioms of arithmetic (natural numbers)

Peano Arithmetic

- Axioms of arithmetic (natural numbers) – self-reference.

Peano Arithmetic

- Axioms of arithmetic (natural numbers) – self-reference.
- $\mathcal{L}_{\mathsf{a}} = \{<, +, \cdot, s, 0\}$

Peano Arithmetic

- Axioms of arithmetic (natural numbers) – self-reference.
- $\mathcal{L}_{\mathsf{a}} = \{<, +, \cdot, s, 0\}$
- Gödel's diagonal Lemma and the incompleteness theorems.

Axioms of Peano Arithmetic (PA)

- $\forall x(s(x) \neq 0)$
- $\forall x, y(s(x) = s(y) \rightarrow x = y)$
- $\forall x(x + 0 = x)$
- $\forall x, y(x + s(y)) = s(x + y)$
- $\forall x(x \cdot 0 = 0)$
- $\forall x, y(x \cdot s(y) = (x \cdot y) + x)$
- $\forall x(\neg x < 0)$
- $\forall x, y(x < s(y) \leftrightarrow (x < y \vee x = y))$
- $\forall x(0 < x \leftrightarrow 0 = x)$
- $\forall x, y(s(x) < y \leftrightarrow (x < y \wedge y \neq s(x)))$
- For all formulae $\phi(x)$,

$$\left(\left(\phi(0) \wedge (\forall x(\phi(x) \rightarrow \phi(x+1)))\right) \rightarrow \forall x \phi(x)\right).$$

Models of arithmetic

- Standard model: $\mathcal{N} = \langle \mathbb{N}; <^{\mathcal{N}}; +^{\mathcal{N}}, \cdot^{\mathcal{N}} ; s^{\mathcal{N}}; 0^{\mathcal{N}} \rangle$.

Models of arithmetic

- Standard model: $\mathcal{N} = \langle \mathbb{N}; <^{\mathcal{N}}; +^{\mathcal{N}}, \cdot^{\mathcal{N}}\ ; s^{\mathcal{N}};\ 0^{\mathcal{N}} \rangle$.
- Non-standard models: $\mathfrak{M} = \langle M; <^{\mathfrak{M}}; +^{\mathfrak{M}}, \cdot^{\mathfrak{M}}; s^{\mathfrak{M}}; 0^{\mathfrak{M}} \rangle$.

Models of arithmetic

- Standard model: $\mathcal{N} = \langle \mathbb{N}; <^{\mathcal{N}}; +^{\mathcal{N}}, \cdot^{\mathcal{N}} ; s^{\mathcal{N}}; 0^{\mathcal{N}} \rangle$.
- Non-standard models: $\mathfrak{M} = \langle M; <^{\mathfrak{M}}; +^{\mathfrak{M}}, \cdot^{\mathfrak{M}}; s^{\mathfrak{M}}; 0^{\mathfrak{M}} \rangle$.
- There is a $c \in M$ such that for any $n \in \mathbb{N}$, $n < c$.

Models of arithmetic

- Standard model: $\mathcal{N} = \langle \mathbb{N}; <^{\mathcal{N}}; +^{\mathcal{N}}, \cdot^{\mathcal{N}}\ ; s^{\mathcal{N}}; 0^{\mathcal{N}} \rangle$.
- Non-standard models: $\mathfrak{M} = \langle M; <^{\mathfrak{M}}; +^{\mathfrak{M}}, \cdot^{\mathfrak{M}}; s^{\mathfrak{M}}; 0^{\mathfrak{M}} \rangle$.
- There is a $c \in M$ such that for any $n \in \mathbb{N}$, $n < c$.
- They are not isomorphic to each other.

Gödel coding

- $f : Form_{\mathcal{L}_a} \longrightarrow \mathbb{N}$

Gödel coding

- $f : Form_{\mathcal{L}_a} \longrightarrow \mathbb{N}$
- f is a recursive function

Gödel coding

- $f : Form_{\mathcal{L}_a} \longrightarrow \mathbb{N}$
- f is a recursive function
- $im(f) \subseteq \mathbb{N}$ is recursive

Gödel coding

- $f : Form_{\mathcal{L}_a} \longrightarrow \mathbb{N}$
- f is a recursive function
- $im(f) \subseteq \mathbb{N}$ is recursive
- f^{-1} is a recursive function

Gödel coding

—

$\mathcal{L}_a$-symbols	Natural numbers
0	0
s	1
$+$	2
$\cdot$	3
$<$	4
$=$	5
$\wedge$	6
$\vee$	7
$\neg$	8
$\exists$	9
$\forall$	10
v_i	$(11, i)$

Theorems

Theorem (Diagonal Lemma)

For any formula $\phi(x)$*, there is a sentence* θ *such that*

$$\mathsf{PA} \vdash \phi(\ulcorner \theta \urcorner) \leftrightarrow \theta.$$

Theorems

Theorem (Diagonal Lemma)

For any formula $\phi(x)$, there is a sentence θ such that

$$\mathsf{PA} \vdash \phi(\ulcorner \theta \urcorner) \leftrightarrow \theta.$$

Proof

We define a function diag $: \mathbb{N} \to \mathbb{N}$ *in the following way:*

$$diag(n) = \begin{cases} \ulcorner \forall y (y = n \to \sigma(y)) \urcorner, & \textit{if } n = \langle \sigma(x) \rangle \textit{ for some formul} \\ 0, & \textit{otherwise.} \end{cases}$$

Definition (Provability predicate)

For any $\mathcal{L}_a$ formula ϕ, the provability predicate in PA *is defined in the following way.*

- *If* $\mathsf{PA} \vdash \phi$ *then* $\mathsf{PA} \vdash Prv(\ulcorner \phi \urcorner)$
- $\mathsf{PA} \vdash Prv(\ulcorner \phi \rightarrow \psi \urcorner) \rightarrow (Prv(\ulcorner \phi \urcorner) \rightarrow Prv(\ulcorner \psi \urcorner))$
- $\mathsf{PA} \vdash Prv(\ulcorner \phi \urcorner) \rightarrow Prv(\ulcorner Prv(\ulcorner \phi \urcorner) \urcorner)$

Theorems

Theorem (Gödel's first incompleteness theorem)

There is a sentence G such that

$$\mathsf{PA} \vdash G \leftrightarrow \neg Prv(\ulcorner G \urcorner).$$

Theorems

Theorem (Gödel's first incompleteness theorem)

There is a sentence G such that

$$\mathsf{PA} \vdash G \leftrightarrow \neg Prv(\ulcorner G \urcorner).$$

Theorem (Gödel's second incompleteness theorem)

Assume PA *is consistent and let* $Con_{\mathsf{PA}} := \neg Prv(\ulcorner \bot \urcorner)$ *be the sentence defining the consistency of* PA. *Then*

$$\mathsf{PA} \nvdash Con_{\mathsf{PA}}$$

i.e. PA *cannot prove its own consistency.*

Liar sentence

Theorem

Assume the Tarski-biconditionals for all sentences in PA. *Let* T *be a predicate defining truth in* PA. T *is undefinable in* $\mathcal{L}_a$.

Liar sentence

Theorem

Assume the Tarski-biconditionals for all sentences in PA*. Let T be a predicate defining truth in* PA*. T is undefinable in $\mathcal{L}_a$.*

Proof

Assume for a contradiction that T is a definable in $\mathcal{L}_a$. Then by the Diagonal Lemma, there is a sentence θ such that

$$\mathsf{PA} \vdash \theta \leftrightarrow \neg T(\ulcorner \theta \urcorner).$$

Then by soundness, $\mathcal{N} \vDash \theta \leftrightarrow \neg \mathcal{T}(\ulcorner \theta \urcorner)$. But by the TB, $\mathcal{N} \vDash \theta \leftrightarrow \mathcal{T}(\ulcorner \theta \urcorner)$.

Typed theories of truth

- We cannot assert truth over a sentence containing truth.

Typed theories of truth

- We cannot assert truth over a sentence containing truth.
- It is true$_1$ that it is true$_0$ that snow is white.

Typed theories of truth

- We cannot assert truth over a sentence containing truth.
- It is true$_1$ that it is true$_0$ that snow is white.
- We don't have a problem with the Liar sentence anymore.

Typed theories of truth

- We cannot assert truth over a sentence containing truth.
- It is true$_1$ that it is true$_0$ that snow is white.
- We don't have a problem with the Liar sentence anymore.
- Assert truth over sentences in PA.

The compositional theory of truth

- $\forall s \forall t (T(s = t) \leftrightarrow val(s) = val(t))$
- $\forall x (sent(x) \rightarrow (T(\neg x) \leftrightarrow \neg T(x)))$
- $\forall x \forall y (sent(x \wedge y) \rightarrow (T(x \wedge y) \leftrightarrow T(x) \wedge T(y)))$
- $\forall x \forall y (sent(x \vee y) \rightarrow (T(x \vee y) \leftrightarrow T(x) \vee T(y)))$
- $\forall v \forall x (sent(\forall vx) \rightarrow (T(\forall vx) \leftrightarrow \forall t T(x(t/v))))$
- $\forall v \forall x (sent(\exists vx) \rightarrow (T(\exists vx) \leftrightarrow \exists t T(x(t/v))))$

Proof theoretic and model theoretic conservativities

- Tarski biconditionals are valid in CT

Proof theoretic and model theoretic conservativities

- Tarski biconditionals are valid in CT
- CT is neither proof theoretically nor model theoretically conservative. It can prove the consistency of PA.

Proof theoretic and model theoretic conservativities

- Tarski biconditionals are valid in CT
- CT is neither proof theoretically nor model theoretically conservative. It can prove the consistency of PA.
- Solution: Restrict the induction axiom schema from PA, to get CT^{-}.

Proof theoretic and model theoretic conservativities

- Tarski biconditionals are valid in CT
- CT is neither proof theoretically nor model theoretically conservative. It can prove the consistency of PA.
- Solution: Restrict the induction axiom schema from PA, to get CT^{-}.
- CT^{-} is proof theoretically conservative. (Enayat & Visser (2013) and Leigh (2013))

Proof theoretic and model theoretic conservativities

- Tarski biconditionals are valid in CT
- CT is neither proof theoretically nor model theoretically conservative. It can prove the consistency of PA.
- Solution: Restrict the induction axiom schema from PA, to get CT^-.
- CT^- is proof theoretically conservative. (Enayat & Visser (2013) and Leigh (2013))
- Tarski biconditionals are valid in CT^-

Satisfaction classes

Definition (Satisfaction class)

$S \subseteq \mathbb{N}^2$ is a *satisfaction class* of a model $\mathfrak{M}$ if

$$S = \{(\ulcorner \phi(x) \urcorner, c) | \mathfrak{M} \models \phi(c)\}$$

- $S_{\mathfrak{M}}(\ulcorner \phi \urcorner, c) \Leftrightarrow \mathfrak{M} \models T(\ulcorner \phi(c) \urcorner)$

Satisfaction classes

Definition (Satisfaction class)

$S \subseteq \mathbb{N}^2$ is a *satisfaction class* of a model $\mathfrak{M}$ if

$$S = \{(\ulcorner \phi(x) \urcorner, c) | \mathfrak{M} \models \phi(c)\}$$

- $S_{\mathfrak{M}}(\ulcorner \phi \urcorner, c) \Leftrightarrow \mathfrak{M} \models T(\ulcorner \phi(c) \urcorner)$
- We expand a model $\mathfrak{M} \models \mathrm{PA}$ by adding the satisfaction class to $\mathfrak{M}$ to get a model of CT^-.

Proof theoretic and model theoretic conservativities

- Is CT^- model theoretically conservative?

Proof theoretic and model theoretic conservativities

- Is CT^- model theoretically conservative? No (By Lachlan's theorem)

Proof theoretic and model theoretic conservativities

- Is CT^- model theoretically conservative? No (By Lachlan's theorem)
- There are non-standard models of CT^- extending PA

Proof theoretic and model theoretic conservativities

- Is CT^- model theoretically conservative? No (By Lachlan's theorem)
- There are non-standard models of CT^- extending PA such that

$$\mathfrak{M} \models T(\ulcorner (0=1) \vee \ldots \vee (0=1) \urcorner).$$

(Kotlarski, Krajewski, Lachlans (1981))

Proof theoretic and model theoretic conservativities

- Is CT^- model theoretically conservative? No (By Lachlan's theorem)
- There are non-standard models of CT^- extending PA such that

  $$\mathfrak{M} \models T(\ulcorner (0 = 1) \vee \ldots \vee (0 = 1) \urcorner).$$

  (Kotlarski, Krajewski, Lachlans (1981))
- Call the satisfaction class S that contains arithmetically false sentences to be *pathological*.

Solutions?

- Eliminate pathological satisfaction classes, containing arithmetically false sentences.

Solutions?

- Eliminate pathological satisfaction classes, containing arithmetically false sentences.
- Cieslinski (2011) adds sentences such as $Prov_{\mathrm{PA}}(x) \rightarrow T(x)$.

Solutions?

- Eliminate pathological satisfaction classes, containing arithmetically false sentences.
- Cieslinski (2011) adds sentences such as $Prov_{\mathrm{PA}}(x) \rightarrow T(x)$.
- All these theories are not proof theoretically conservative.

Solutions?

- Eliminate pathological satisfaction classes, containing arithmetically false sentences.
- Cieslinski (2011) adds sentences such as $Prov_{\mathrm{PA}}(x) \to T(x)$.
- All these theories are not proof theoretically conservative.
- What can we do?

Solutions?

- Tarski's model theoretic definition of truth?

Solutions?

- Tarski's model theoretic definition of truth?
- Non-standard models give a non-standard interpretation?

Solutions?

- Tarski's model theoretic definition of truth?
- Non-standard models give a non-standard interpretation?
- Second order arithmetic with full-semantics?

Solutions?

- Tarski's model theoretic definition of truth?
- Non-standard models give a non-standard interpretation?
- Second order arithmetic with full-semantics?
- Is model theoretic conservativity for deflationists?

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.
- PA cannot define truth.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.
- PA cannot define truth.
- We add the compositional axioms and T predicate to attain CT.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.
- PA cannot define truth.
- We add the compositional axioms and T predicate to attain CT.
- CT fails to be proof/model theoretically conservative. So we restrict it to CT^-.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.
- PA cannot define truth.
- We add the compositional axioms and T predicate to attain CT.
- CT fails to be proof/model theoretically conservative. So we restrict it to CT$^-$.
- CT$^-$ is not model theoretically conservative, and it states arithmetically false sentences are true.

Summary and conclusion

- Deflationists desire proof theoretic and model theoretic conservativities.
- We start with PA as our base theory as it allows self-reference.
- PA cannot define truth.
- We add the compositional axioms and T predicate to attain CT.
- CT fails to be proof/model theoretically conservative. So we restrict it to CT^-.
- CT^- is not model theoretically conservative, and it states arithmetically false sentences are true.
- Cieslinski's elimination methods fails to save CT^-.