A proof-theoretic account of classical principles of truth

Graham E. Leigh

University of Oxford

Numbers and Truth conference,
Göteborg, 19th October 2012

Outline

Framework

Consistency

- Classical vs. intuitionistic logic
- Non-trivial models of $A \to \mathrm{T}A$

Proof theoretic strength

- Truth over classical logic
- Truth over intuitionistic logic
- A surprising result?

Outline

Framework

Consistency
- Classical vs. intuitionistic logic
- Non-trivial models of $A \to \mathrm{T}A$

Proof theoretic strength
- Truth over classical logic
- Truth over intuitionistic logic
- A surprising result?

Framework

$\mathcal{L}_{\mathrm{T}}$ denotes the language of arithmetic expanded by a fresh predicate symbol T.

$\mathrm{T}(A)$ represents "A is true."

PA_{T} is Peano arithmetic in the language $\mathcal{L}_{\mathrm{T}}$;
HA_{T} is Heyting arithmetic in $\mathcal{L}_{\mathrm{T}}$.

Definition

$\mathrm{Base}^i_{\mathrm{T}}$ is the theory extending HA_{T} by the following axioms:

1. $\forall A \forall B(\mathrm{T}A \wedge \mathrm{T}(A \to B) \to \mathrm{T}B)$.
2. $\forall x(\mathit{valid}^i(x) \to \mathrm{T}(x))$.
3. $\mathrm{T}A$, if A is the universal closure of an axiom of PRA.

$\mathrm{Base}_{\mathrm{T}}$ is $\mathrm{Base}^i_{\mathrm{T}} + \forall A\, \mathrm{T}(A \vee \neg A) + \mathrm{LEM}$.

Axioms and rules

Axiom Schemata

$A \to \mathrm{T}A$ $\qquad$ $\mathrm{T}A \to A$

Axioms and rules

Axiom Schemata

$$A \to \mathrm{T}A \qquad\qquad \mathrm{T}A \to A$$

Axioms

$$\mathrm{T}A \to \mathrm{TT}A \qquad\qquad \mathrm{TT}A \to \mathrm{T}A$$

$$\text{(Comp)} \quad \mathrm{T}A \vee \mathrm{T}\neg A \qquad\qquad \neg(\mathrm{T}A \wedge \mathrm{T}\neg A) \quad \text{(Cons)}$$

Axioms and rules

Axiom Schemata

$$A \to \mathrm{T}A \qquad\qquad \mathrm{T}A \to A$$

Axioms

$$\mathrm{T}A \to \mathrm{TT}A \qquad\qquad \mathrm{TT}A \to \mathrm{T}A$$

$$\text{(Comp)}\quad \mathrm{T}A \lor \mathrm{T}\neg A \qquad\qquad \neg(\mathrm{T}A \land \mathrm{T}\neg A)\quad \text{(Cons)}$$

$$\text{(Comp(w))}\quad \neg\mathrm{T}A \to \mathrm{T}\neg A \qquad\qquad \forall x \mathrm{T}A(\dot{x}) \to \mathrm{T}\forall x A(x)\quad (\forall\text{-Inf})$$

$$(\lor\text{-Inf})\quad \mathrm{T}(A \lor B) \to \mathrm{T}A \lor \mathrm{T}B \qquad\qquad \mathrm{T}\exists x A(x) \to \exists x \mathrm{T}A(\dot{x})\quad (\exists\text{-Inf})$$

$$(\mathrm{T}A \to \mathrm{T}B) \to \mathrm{T}(A \to B)$$

Axioms and rules

Axiom Schemata

$$A \rightarrow \mathrm{T}A \qquad\qquad \mathrm{T}A \rightarrow A$$

Axioms

$$\mathrm{T}A \rightarrow \mathrm{TT}A \qquad\qquad \mathrm{TT}A \rightarrow \mathrm{T}A$$

$$\text{(Comp)} \quad \mathrm{T}A \vee \mathrm{T}\neg A \qquad\qquad \neg(\mathrm{T}A \wedge \mathrm{T}\neg A) \quad \text{(Cons)}$$

$$\text{(Comp(w))} \quad \neg\mathrm{T}A \rightarrow \mathrm{T}\neg A \qquad\qquad \forall x \mathrm{T}A(\dot{x}) \rightarrow \mathrm{T}\forall x A(x) \quad (\forall\text{-Inf})$$

$$(\vee\text{-Inf}) \quad \mathrm{T}(A \vee B) \rightarrow \mathrm{T}A \vee \mathrm{T}B \qquad\qquad \mathrm{T}\exists x A(x) \rightarrow \exists x \mathrm{T}A(\dot{x}) \quad (\exists\text{-Inf})$$

$$(\mathrm{T}A \rightarrow \mathrm{T}B) \rightarrow \mathrm{T}(A \rightarrow B)$$

Rules of Inference

$$A / \mathrm{T}A \qquad\qquad \neg A / \neg\mathrm{T}A$$

$$\mathrm{T}A / A \qquad\qquad \neg\mathrm{T}A / \neg A$$

Axioms and rules

Axiom Schemata

$$A \to \mathrm{T}A \qquad\qquad \mathrm{T}A \to A$$

Axioms

$$\mathrm{T}A \to \mathrm{TT}A \qquad\qquad \mathrm{TT}A \to \mathrm{T}A$$

$$\text{(Comp)} \quad \mathrm{T}A \vee \mathrm{T}\neg A \qquad\qquad \neg(\mathrm{T}A \wedge \mathrm{T}\neg A) \quad \text{(Cons)}$$

$$\text{(Comp(w))} \quad \neg\mathrm{T}A \to \mathrm{T}\neg A \qquad\qquad \forall x \mathrm{T}A(\dot{x}) \to \mathrm{T}\forall x A(x) \quad (\forall\text{-Inf})$$

$$(\vee\text{-Inf}) \quad \mathrm{T}(A \vee B) \to \mathrm{T}A \vee \mathrm{T}B \qquad\qquad \mathrm{T}\exists x A(x) \to \exists x \mathrm{T}A(\dot{x}) \quad (\exists\text{-Inf})$$

$$(\mathrm{T}A \to \mathrm{T}B) \to \mathrm{T}(A \to B)$$

Rules of Inference

$$A/\mathrm{T}A \qquad\qquad \neg A/\neg\mathrm{T}A$$

$$\mathrm{T}A/A \qquad\qquad \neg\mathrm{T}A/\neg A$$

A "theory of truth" is a theory $\mathrm{Base}^i_{\mathrm{T}} + S$, or $\mathrm{Base}_{\mathrm{T}} + S$ where S is some subset of the above principles.

Outline

Classical vs. intuitionistic logic

Theorem (Friedman & Sheard '87)

Of the fifteen principles on the previous side, there are exactly nine maximal consistent subsets over $\mathrm{Base_T}$.

Classical vs. intuitionistic logic

Theorem (Friedman & Sheard '87)

Of the fifteen principles on the previous side, there are exactly nine maximal consistent subsets over $\mathrm{Base}_{\mathrm{T}}$.

Over $\mathrm{Base}_{\mathrm{T}}$ the following principles are equivalent.

T-Comp:	$\mathrm{T}A \vee \mathrm{T}\neg A$
T-Comp(w):	$\neg \mathrm{T}A \rightarrow \mathrm{T}\neg A$
$\vee$-Inf:	$\mathrm{T}(A \vee B) \rightarrow \mathrm{T}A \vee \mathrm{T}B$

Over intuitionistic logic, each pair of the above can be separated.

Classical vs. intuitionistic logic

Theorem (Friedman & Sheard '87)

Of the fifteen principles on the previous side, there are exactly nine maximal consistent subsets over $\mathrm{Base_T}$.

Over $\mathrm{Base_T}$ the following principles are equivalent.

T-Comp:	$\mathrm{T}A \vee \mathrm{T}\neg A$
T-Comp(w):	$\neg \mathrm{T}A \rightarrow \mathrm{T}\neg A$
$\vee$-Inf:	$\mathrm{T}(A \vee B) \rightarrow \mathrm{T}A \vee \mathrm{T}B$

Over intuitionistic logic, each pair of the above can be separated. Still:

Theorem (L & Rathjen '12)

Over $\mathrm{Base}^{i}_{\mathrm{T}}$ *there are exactly nine maximal consistent sets.*

Non-trivial models of $A \to \mathrm{T}A$

The pair of principles $A \to \mathrm{T}A$ and $\mathrm{T}A/A$ are inconsistent over $\mathrm{Base}_{\mathrm{T}}$.

Theorem

The following collection of axioms is consistent with $\mathrm{Base}^{i}_{\mathrm{T}}$.

1) $A \to \mathrm{T}A$	*4)* $\mathrm{T}A/A$
2) $\mathrm{T}(A \vee B) \to \mathrm{T}A \vee \mathrm{T}B$	*5)* $(\mathrm{T}A \to \mathrm{T}B) \to \mathrm{T}(A \to B)$
3) $\forall x \mathrm{T}[A(\dot{x})] \to \mathrm{T}[\forall x A(x)]$	*6)* $\mathrm{T}[\exists x A(x)] \to \exists x \mathrm{T}[A(\dot{x})]$

Non-trivial models of $A \to \mathrm{T}A$

The pair of principles $A \to \mathrm{T}A$ and $\mathrm{T}A/A$ are inconsistent over Base_T.

Theorem

The following collection of axioms is consistent with $\mathrm{Base}^i_\mathrm{T}$.

1) $A \to \mathrm{T}A$
2) $\mathrm{T}(A \vee B) \to \mathrm{T}A \vee \mathrm{T}B$
3) $\forall x \mathrm{T}[A(\dot{x})] \to \mathrm{T}[\forall x A(x)]$
4) $\mathrm{T}A/A$
5) $(\mathrm{T}A \to \mathrm{T}B) \to \mathrm{T}(A \to B)$
6) $\mathrm{T}[\exists x A(x)] \to \exists x \mathrm{T}[A(\dot{x})]$

Proof. Apply the rule-of-revision with intuitionistic Kripke models. Define a hierarchy of models such that $\mathfrak{A}_0 \models \forall A\, \mathrm{T}A$ and

$$(\mathfrak{A}_{n+1} \models \mathrm{T}A \Leftrightarrow \mathfrak{A}_n \models A) \quad \text{and} \quad (\mathfrak{A}_{n+1} \models A \Rightarrow \mathfrak{A}_n \models A).$$

Then $\mathrm{Base}^i_\mathrm{T} + 1\text{–}6 \vdash A$ implies $\mathfrak{A}_n \models A$ for every n.

Outline

Framework

Consistency

- Classical vs. intuitionistic logic
- Non-trivial models of $A \to \mathrm{T}A$

Proof theoretic strength

- Truth over classical logic
- Truth over intuitionistic logic
- A surprising result?

Truth over classical logic

Over Base_{T}, a proof-theoretic analysis of the maximal consistent theories yields:

Maximal consistent set	Equivalent theories
$(A \to \mathrm{T}A) + \forall\text{-Inf} + \dots$	PA
$\text{Cons} + \text{Comp} + (\mathrm{T}A \to \mathrm{TT}A) + \forall\text{-Inf} + \dots$	ACA; $\mathrm{PA} + \mathsf{TI}(<\epsilon_{\epsilon_0})$
$\text{Cons} + \text{Comp} + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \dots$	ACA; $\mathrm{PA} + \mathsf{TI}(<\epsilon_{\epsilon_0})$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + \text{Cons} + \text{Comp} + \forall\text{-Inf} + \dots$	ACA_0^+; $\mathrm{RA}_{<\omega}$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{T}A \to \mathrm{TT}A) + \forall\text{-Inf} + \dots$	ACA_0^+; $\mathrm{RA}_{<\omega}$
$(\mathrm{T}A/A) + (\mathrm{T}A \to \mathrm{TT}A) + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \dots$	ACA_0^+; $\mathrm{RA}_{<\omega}$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \dots$	$\Sigma^1_1\text{-DC}_0$; ID_1^*; $\mathrm{PA} + \mathsf{TI}(<\varphi_\omega 0)$
$(\mathrm{T}A \to A) + \forall\text{-Inf} + \dots;$	ID_1; $\mathrm{KP}\omega$

(Cantini '90; Halbach '94; L & Rathjen '10)

Truth over intuitionistic logic

Over $\mathrm{Base_T}$, a proof-theoretic analysis of the maximal consistent theories yields:

Maximal consistent set	Equivalent theories
$(A \to \mathrm{T}A) + \forall\text{-Inf} + \ldots$	HA
$\mathrm{Cons} + \mathrm{Comp} + (\mathrm{T}A \to \mathrm{TT}A) + \forall\text{-Inf} + \ldots$	ACA; $\mathrm{PA} + \mathsf{TI}(<\epsilon_{\epsilon_0})$
$\mathrm{Cons} + \mathrm{Comp} + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \ldots$	ACA; $\mathrm{PA} + \mathsf{TI}(<\epsilon_{\epsilon_0})$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + \mathrm{Cons} + \mathrm{Comp} + \forall\text{-Inf} + \ldots$	ACA_0^+; $\mathrm{RA}_{<\omega}$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{T}A \to \mathrm{TT}A) + (A \to \mathrm{T}A) + \ldots$	ACA_0^{i+}; $\mathrm{RA}^i_{<\omega}$
$(\mathrm{T}A/A) + (\mathrm{T}A \to \mathrm{TT}A) + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \ldots$	ACA_0^{i+}; $\mathrm{RA}^i_{<\omega}$
$(\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \ldots$	$\Sigma^1_1\text{-}\mathrm{DC}^i_0$; ~~$\mathrm{ID}_1^{i*}$;~~ $\mathrm{HA} + \mathsf{TI}(<\varphi_\omega 0)$
$(\mathrm{T}A \to A) + \forall\text{-Inf} + \ldots;$	ID^i_1; $\mathrm{KP}\omega^i$

(L '12)

A surprising result?

Consider the theory

$$\mathcal{F}^i: \mathrm{Base}^i_{\mathrm{T}} + (\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{TT}A \rightarrow \mathrm{T}A) + \forall\text{-Inf} + \\ + \mathrm{Comp(w)} + \exists\text{-Inf} + \vee\text{-Inf}.$$

A surprising result?

Consider the theory

$$\mathcal{F}^i\colon \mathrm{Base}^i_{\mathrm{T}} + (\mathrm{T}A/A) + (A/\mathrm{T}A) + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \\ + \mathrm{Comp(w)} + \exists\text{-Inf} + \vee\text{-Inf}.$$

Let

$$\mathrm{Th}_0 = \mathrm{Base}^i_{\mathrm{T}} + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \dots$$
$$\mathrm{Th}_{m+1} = \mathrm{Base}^i_{\mathrm{T}} + (\mathrm{TT}A \to \mathrm{T}A) + \forall\text{-Inf} + \dots + \{\mathrm{T}A \mid \mathrm{Th}_m \vdash A\}.$$

Then

1. Th_m has the disjunction and existence property for each m;
2. $\mathrm{Th}_m \vdash_\omega \mathrm{T}A$ implies $\mathrm{Th}_m \vdash_\omega A$.

So

$$\mathcal{F}^i \vdash \mathrm{T}A \quad \Rightarrow \quad \exists m\colon \mathrm{Th}_m \vdash_\omega A,$$

suggesting $\mathcal{F}^i \hookrightarrow \mathrm{ID}^{*i}_1$.

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

1. Express derivability in Th_m by a Gentzen-style sequent calculus: $\Gamma \Rightarrow^m_\alpha A$.

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

1. Express derivability in Th_m by a Gentzen-style sequent calculus: $\Gamma \Rightarrow^m_\alpha A$.
2. Aim to prove that if $\varnothing \Rightarrow^m_\alpha \mathrm{T}\ulcorner A \urcorner$ then $\varnothing \Rightarrow^m_{f_m(\alpha)} A$.

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

1. Express derivability in Th_m by a Gentzen-style sequent calculus: $\Gamma \Rightarrow^m_\alpha A$.
2. Aim to prove that if $\varnothing \Rightarrow^m_\alpha \mathrm{T}\ulcorner A \urcorner$ then $\varnothing \Rightarrow^m_{f_m(\alpha)} A$.
3. For ordinals α, β, define $A^{(\alpha,\beta)}$ by:
   $\mathrm{T}(s)^{(\alpha,\beta)}$ iff $\varnothing \Rightarrow^m_{\varphi_m \beta} s$;
   $(A \wedge B)^{(\alpha,\beta)}$ iff $A^{(\alpha,\beta)}$ and $B^{(\alpha,\beta)}$, $\ldots$,
   $(A \to B)^{(\alpha,\beta)}$ iff $A^{(\beta,\alpha)}$ implies $B^{(\alpha,\beta)}$.

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

1. Express derivability in Th_m by a Gentzen-style sequent calculus: $\Gamma \Rightarrow^m_\alpha A$.
2. Aim to prove that if $\varnothing \Rightarrow^m_\alpha \mathrm{T}\ulcorner A \urcorner$ then $\varnothing \Rightarrow^m_{f_m(\alpha)} A$.
3. For ordinals α, β, define $A^{(\alpha,\beta)}$ by:
   $\mathrm{T}(s)^{(\alpha,\beta)}$ iff $\varnothing \Rightarrow^m_{\varphi_m \beta} s$;
   $(A \wedge B)^{(\alpha,\beta)}$ iff $A^{(\alpha,\beta)}$ and $B^{(\alpha,\beta)}$, $\ldots$,
   $(A \rightarrow B)^{(\alpha,\beta)}$ iff $A^{(\beta,\alpha)}$ implies $B^{(\alpha,\beta)}$.
4. Soundness of m-derivability:

$$\text{If } \Gamma \Rightarrow^m_\alpha A \text{ then } (\bigwedge \Gamma \rightarrow A)^{(\gamma,\gamma+\omega^\alpha)} \text{ for every } \gamma.$$

A surprising result?

How do we formally prove $\mathrm{Th}_m \vdash_\omega \mathrm{T}A \Rightarrow \mathrm{Th}_m \vdash_\omega A$?

1. Express derivability in Th_m by a Gentzen-style sequent calculus: $\Gamma \Rightarrow^m_\alpha A$.
2. Aim to prove that if $\varnothing \Rightarrow^m_\alpha \mathrm{T}\ulcorner A \urcorner$ then $\varnothing \Rightarrow^m_{f_m(\alpha)} A$.
3. For ordinals α, β, define $A^{(\alpha,\beta)}$ by:
   $\mathrm{T}(s)^{(\alpha,\beta)}$ iff $\varnothing \Rightarrow^m_{\varphi_m\beta} s$;
   $(A \wedge B)^{(\alpha,\beta)}$ iff $A^{(\alpha,\beta)}$ and $B^{(\alpha,\beta)}$, ...,
   $(A \rightarrow B)^{(\alpha,\beta)}$ iff $A^{(\beta,\alpha)}$ implies $B^{(\alpha,\beta)}$.
4. Soundness of m-derivability:

$$\text{If } \Gamma \Rightarrow^m_\alpha A \text{ then } (\bigwedge \Gamma \rightarrow A)^{(\gamma,\gamma+\omega^\alpha)} \text{ for every } \gamma.$$

Proposition 1. If $\mathcal{F}^i \vdash_m A$ then $\varnothing \Rightarrow^m_{\varphi_{m+1}0} A$.
Proposition 2. If $\mathcal{F}^i \vdash_m A \in \mathcal{L}$ then $\mathrm{HA} + \mathrm{TI}(<\varphi_{m+2}0) \vdash A$.