

The Inconsistency in Gödel's Ontological Argument — A Success Story for AI in Metaphysics —

Christoph Benzmueller¹ and Bruno Woltzenlogel Paleo

¹Supported by the German Research Foundation (grant BE2501/12-1,2)

Motivation

Vision of Leibniz (1646–1716): *Calculemus!*



Quo facto, quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo . . . dicere: calculemus. (Leibniz, 1684)



Required:
characteristica universalis and **calculus ratiocinator**

If controversies were to arise, there would be no more need of disputation between two philosophers than between two accountants. For it would suffice to take their pencils in their hands, to sit down to their slates, and to say to each other . . . : Let us calculate.

(Translation by Russell)

Scott's and Gödel's Versions

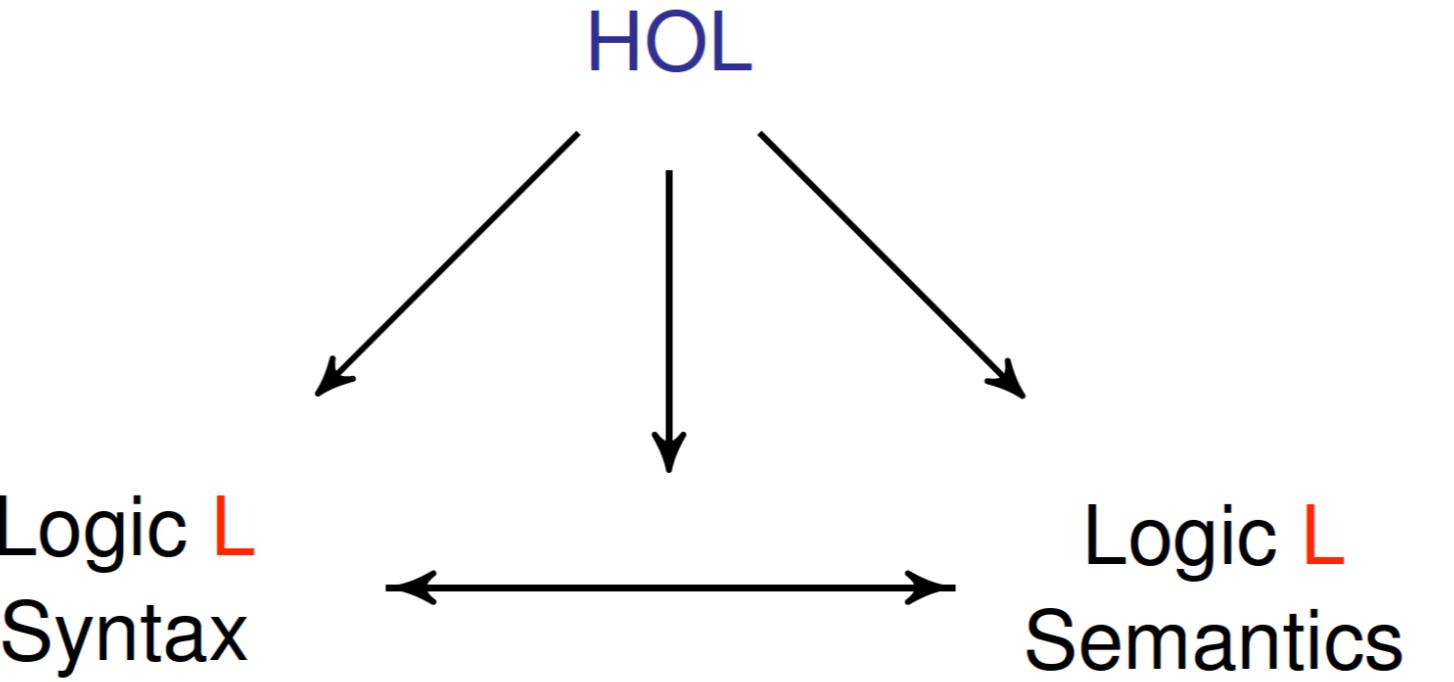
Scott's Version of Gödel's Axioms, Definitions and Theorems

Axiom A1 Either a property or its negation is positive, but not both:	$\forall\phi[P(\neg\phi) \leftrightarrow \neg P(\phi)]$
Axiom A2 A property necessarily implied by a positive property is positive:	$\forall\phi\psi[(P(\phi) \wedge \Box\forall x[\phi(x) \rightarrow \psi(x)]) \rightarrow P(\psi)]$
Thm. T1 Positive properties are possibly exemplified:	$\forall\phi[P(\phi) \rightarrow \Diamond\exists x\phi(x)]$
Def. D1 A God-like being possesses all positive properties:	$G(x) \leftrightarrow \forall\phi[P(\phi) \rightarrow \phi(x)]$
Axiom A3 The property of being God-like is positive:	$P(G)$
Cor. C Possibly, God exists:	$\Diamond\exists xG(x)$
Axiom A4 Positive properties are necessarily positive:	$\forall\phi[P(\phi) \rightarrow \Box P(\phi)]$
Def. D2 An essence of an individual is a property possessed by it and necessarily implying any of its properties:	$\phi \text{ ess. } x \leftrightarrow \phi(x) \wedge \forall\psi(\psi(x) \rightarrow \Box\forall y(\phi(y) \rightarrow \psi(y)))$
Thm. T2 Being God-like is an essence of any God-like being:	$\forall x[G(x) \rightarrow G \text{ ess. } x]$
Def. D3 Necessary existence of an individual is the necessary exemplification of all its essences:	$NE(x) \leftrightarrow \forall\theta[\phi \text{ ess. } x \rightarrow \Box\forall y\phi(y)]$
Axiom A5 Necessary existence is a positive property:	$P(NE)$
Thm. T3 Necessarily, God exists:	$\Box\exists xG(x)$

Difference to Gödel (who omits this conjunct)

Approach: Semantic Embedding

HOL as a Universal (Meta-)Logic via Semantic Embeddings



Examples for L we have already studied:

Modal Logics, Conditional Logics, Intuitionistic Logics, Access Control Logics, Nominal Logics, Multivalued Logics (SIXTEEN), Logics based on Neighborhood Semantics, (Mathematical) Fuzzy Logics, Paraconsistent Logics, Free Logic ...

Works also for (first-order & higher-order) quantifiers

Standard Translation of Modal Logic

Embedding HOL in HOL

Example

HOL formula
HOL formula in HOL
expansion
 $\beta\eta$ -normalisation
expansion
 $\beta\eta$ -normalisation
syntactic sugar
expansion
 $\beta\eta$ -normalisation

$\forall w_\mu(\Box\varphi_\mu \rightarrow \Diamond\exists xG(x))_{\mu \rightarrow o}$

$\forall w_\mu(\Box\exists xG(x))_{\mu \rightarrow o}$

$\forall w_\mu(\exists x_\mu(rwu \wedge \Box xG(x))_{\mu \rightarrow o})$

$\forall w_\mu(\exists x_\mu(rwu \wedge (\exists xG(x))_{\mu \rightarrow o}))$

$\forall w_\mu(\exists x_\mu(rwu \wedge (\lambda h_{\gamma \rightarrow \mu \rightarrow o} \lambda w_\mu \exists d_\gamma(hdw)(\exists xG(x))_{\mu \rightarrow o}))$

$\forall w_\mu(\exists x_\mu(rwu \wedge \exists xGx))$

Expansion: user or prover may flexibly choose expansion depth

What are we doing?

In order to prove that φ is valid in HOL,

→ we instead prove that $\text{valid } \varphi_{\mu \rightarrow o}$ can be derived from Ax in HOL.

This can be done with interactive or automated HOL theorem provers.

Modal Logic in Isabelle

```
abbreviation mnot :: "σ⇒σ" ("_¬[_]" [52] 53)
  where "¬φ ≡ λx. ¬φ(x)" 
abbreviation mand :: "σ⇒σ⇒σ" (infixr "Λ" 51)
  where "φΛψ ≡ λx. φ(x)Λψ(x)" 
abbreviation mor :: "σ⇒σ⇒σ" (infixr "V" 50)
  where "φVψ ≡ λx. φ(x)Vψ(x)" 
abbreviation mimp :: "σ⇒σ⇒σ" (infixr "→" 49)
  where "φ→ψ ≡ λx. φ(x)→ψ(x)" 
abbreviation mequ :: "σ⇒σ⇒σ" (infixr "↔" 48)
  where "φ↔ψ ≡ λx. φ(x)↔ψ(x)" 
abbreviation mall :: "(" → "σ)⇒σ" ("V")
  where "Vθ = λx. θ(x)Vx" 
abbreviation mallB :: "(" → "σ)⇒σ" (binder "V" [8] 9)
  where "Vx. θ(x) ≡ Vxθ" 
abbreviation mexi :: "(" → "σ)⇒σ" ("E")
  where "Eθ = λx. θ(x)E(x)" 
abbreviation mexiB :: "(" → "σ)⇒σ" (binder "E" [8] 9)
  where "Ex. θ(x) ≡ Exθ" 
abbreviation mem :: "(" → "σ)⇒σ" (infixr "≡" 52) -- "Equality"
  where "xy ≡ λx. λy. x=y" 
abbreviation meql :: "μ⇒μ⇒σ" (infixr "≡" 52) -- "Leibniz Equality"
  where "xy = Yφ. φ(x)→φ(y)" 
abbreviation mbox :: "σ⇒σ" ("□" [52] 53)
  where "□φ ≡ λx. ∀v. v=x → φ(v)" 
abbreviation mdia :: "σ⇒σ" ("◊" [52] 53)
  where "◊φ ≡ λw. ∃v. w=v ∧ φ(v)"
```

Scott's Version in Isabelle

```
begin consts P :: "(μ=⇒)⇒σ" (* Positive *)
  axiomatization where
    A1: "[|'VΦ. P(Φ) → ¬P(Φ)|] and A2: [|'VΦ. ¬¬P(Φ) → P(Φ)|] and
    definition G where "G(x) = (VΦ. P(Φ) → x)" (* Definition of God *)
  axiomatization where A3: "[|P(G)|] and A4: [|'VΦ. P(Φ) → □P(Φ)|]"
  definition ess (infixr "ess" 85) where
    "Φ ess x = Φ(x) ∧ Φ(x) → □(Φ(y) → Φ(y))" (* Essence *)
  definition NE where "NE(x) = (VΦ. Φ ess x → □(D(Φ)))" (* Necessary Existence *)
  axiomatization where A5: "[|P(NE)|]"
  theorem T3: "|□(E(x))| (* Necessarily, God exists: LEO-II proof in 2,5sec *)
```

sledgehammer [provers = remote_leo2]
by (metis lifting, full_types)

Ala A1 A2 A3 A4 A5 G_def NE_def ess_def

lemma True nitpick [satisfy,user_axioms,expect=genuine] oops
-- (* Consistency is confirmed by Nitpick *)

theorem T3: "|Vx. G(x) → G ess x|"
sledgehammer [provers = remote_leo2]
by (metis AlA A4 G_def ess_def)

lemma Mc: "|'VΦ. Φ → (□Φ)|" -- (* Modal Collapse *)
sledgehammer [provers = remote_satallax, timeout=600]
by (meson T2 T3 ess_def)
end

LEO-II's Refutation of Gödel's Axioms

```
abbreviation mnot :: "σ⇒σ" ("_¬[_]" [52] 53)
  where "¬φ ≡ λx. ¬φ(x)" 
abbreviation mand :: "σ⇒σ⇒σ" (infixr "Λ" 51)
  where "φΛψ ≡ λx. φ(x)Λψ(x)" 
abbreviation mor :: "σ⇒σ⇒σ" (infixr "V" 50)
  where "φVψ ≡ λx. φ(x)Vψ(x)" 
abbreviation mimp :: "σ⇒σ⇒σ" (infixr "→" 49)
  where "φ→ψ ≡ λx. φ(x)→ψ(x)" 
abbreviation mequ :: "σ⇒σ⇒σ" (infixr "↔" 48)
  where "φ↔ψ ≡ λx. φ(x)↔ψ(x)" 
abbreviation mall :: "(" → "σ)⇒σ" ("V")
  where "Vθ = λx. θ(x)Vx" 
abbreviation mallB :: "(" → "σ)⇒σ" (binder "V" [8] 9)
  where "Vx. θ(x) ≡ Vxθ" 
abbreviation mexi :: "(" → "σ)⇒σ" ("E")
  where "Eθ = λx. θ(x)E(x)" 
abbreviation mexiB :: "(" → "σ)⇒σ" (binder "E" [8] 9)
  where "Ex. θ(x) ≡ Exθ" 
abbreviation mem :: "(" → "σ)⇒σ" (infixr "≡" 52) -- "Equality"
  where "xy ≡ λx. λy. x=y" 
abbreviation meql :: "μ⇒μ⇒σ" (infixr "≡" 52) -- "Leibniz Equality"
  where "xy = Yφ. φ(x)→φ(y)" 
abbreviation mbox :: "σ⇒σ" ("□" [52] 53)
  where "□φ ≡ λx. ∀v. v=x → φ(v)" 
abbreviation mdia :: "σ⇒σ" ("◊" [52] 53)
  where "◊φ ≡ λw. ∃v. w=v ∧ φ(v)"
```

Refutation Reconstruction in Isabelle

```
theory GoedelGodWithoutConjunctInEss_K imports QML
begin
consts P :: "(μ=⇒)⇒σ"
definition ess (infixr "ess" 85) where "Φ ess x = (λy. Φ y → □(λy. Φ y → Φ y))"
definition NE where "NE x = ∀y. Φ ess x → □(y=⇒Φ y)"
axiomatization where Ala: "[|'V(λΦ. P(Φ) → □P(Φ))|]" and
  A2: "[|'VΦ. Φ → □P(Φ)|]" and A3: "[|P(G)|]" and A4: "[|'VΦ. P(Φ) → □P(Φ)|]"
definition G where "G(x) = (VΦ. P(Φ) → x)" (* Definition of God *)
definition ess (infixr "ess" 85) where
  "Φ ess x = Φ(x) ∧ Φ(x) → □(Φ(y) → Φ(y))" (* Essence *)
definition NE where "NE(x) = (VΦ. Φ ess x → □(D(Φ)))" (* Necessary Existence *)
axiomatization where A5: "[|P(NE)|]"
theorem T1: "|'V(λΦ. P(Φ) → □P(Φ))|"
sledgehammer [remote_leo2]
by (metis Ala A2)
lemma False
-- sledgehammer [remote_leo2]
by (metis False)
axiomatization where A5: "[|P(NE)|]"
where no the inconsistency follows from A5, Lemmal, NE_def and T1 *
lemma False
-- sledgehammer [remote_leo2]
by (metis False)
end
```

Explaining the Inconsistency

Def. D2* $\phi \text{ ess. } x \leftrightarrow \Diamond\exists x[\psi(\psi(x) \rightarrow \Box\forall y(\phi(y) \rightarrow \psi(y)))]$

Lemma 1 The empty property is an essence of every entity. $\forall x(\Diamond\text{ess. } x)$

Theorem 1 Positive Properties are possibly exemplified. $\forall\phi[P(\phi) \rightarrow \Diamond\exists x\phi(x)]$

Axiom A5

- ▶ by T1, A5: $\Diamond\exists x[NE(x)]$
- ▶ by D3: $\Diamond\exists x[\forall\phi[\phi \text{ ess. } x \rightarrow \Box\forall y(\phi(y) \rightarrow \psi(y))]]$
- ▶ by L1: $\Diamond\exists x[\emptyset \text{ ess. } x \rightarrow \Box\forall y(\emptyset(y) \rightarrow \psi(y))]$
- ▶ by def. of 0: $\Diamond\exists x[T \rightarrow \Box\forall y(T(y) \rightarrow \psi(y))]$
- ▶ by Inconsistency: $\Diamond\exists x[\Box\forall y(T(y) \rightarrow \psi(y)) \wedge \Box\forall y(T(y) \rightarrow \neg\psi(y))]$

Summary of Results

- Inconsistency of Gödel's original axioms has now been verified
- Reason for inconsistency is finally well understood and explained
- Reconstruction of the refutation in Isabelle led to various user interface and performance improvements for the embedding of modal logic in Isabelle
- Higher-order interactive and automated reasoning technology is ready for applications in philosophy