

# **Second Incompleteness Theorem** **(G2)**

**Selmer Bringsjord**

Rensselaer AI & Reasoning (RAIR) Lab  
Department of Cognitive Science  
Department of Computer Science  
Lally School of Management & Technology  
Rensselaer Polytechnic Institute (RPI)  
Troy, New York 12180 USA

4/15/2024



# **Second Incompleteness Theorem** **(G2)**

**Selmer Bringsjord**

Rensselaer AI & Reasoning (RAIR) Lab  
Department of Cognitive Science  
Department of Computer Science  
Lally School of Management & Technology  
Rensselaer Polytechnic Institute (RPI)  
Troy, New York 12180 USA

4/15/2024



**OXFORD**  
UNIVERSITY PRESS

# **Second Incompleteness Theorem** **(G2)**

**Selmer Bringsjord**

Rensselaer AI & Reasoning (RAIR) Lab  
Department of Cognitive Science  
Department of Computer Science  
Lally School of Management & Technology  
Rensselaer Polytechnic Institute (RPI)  
Troy, New York 12180 USA

4/15/2024



**OXFORD**  
UNIVERSITY PRESS



# **Second Incompleteness Theorem** **(G2)**

**Selmer Bringsjord**

Rensselaer AI & Reasoning (RAIR) Lab  
Department of Cognitive Science  
Department of Computer Science  
Lally School of Management & Technology  
Rensselaer Polytechnic Institute (RPI)  
Troy, New York 12180 USA

4/15/2024



**OXFORD**  
UNIVERSITY PRESS



# Gödel's Second Incompleteness Theorem (G2)

**Selmer Bringsjord**

Rensselaer AI & Reasoning (RAIR) Lab  
Department of Cognitive Science  
Department of Computer Science  
Lally School of Management & Technology  
Rensselaer Polytechnic Institute (RPI)  
Troy, New York 12180 USA

4/15/2024

**Note:** This is a version designed for those who have had at least one university-level course in formal logic with coverage through  $\mathcal{L}_1$ .



OXFORD  
UNIVERSITY PRESS



**Background Context ...**

# *Gödel's Great Theorems* (OUP)

by Selmer Bringsjord

- Introduction (“The Wager”)
- Brief Preliminaries (e.g. the propositional calculus & FOL)
- The Completeness Theorem
- The First Incompleteness Theorem
- The Second Incompleteness Theorem
- The Speedup Theorem
- The Continuum-Hypothesis Theorem
- The Time-Travel Theorem
- Gödel’s “God Theorem”
- Could a Finite Machine Match Gödel’s Greatness?



STOP & REVIEW IF NEEDED!

# *Gödel's Great Theorems* (OUP)

by Selmer Bringsjord

- Introduction (“The Wager”)
- Brief Preliminaries (e.g. the propositional calculus & FOL)
- The Completeness Theorem
- The First Incompleteness Theorem
- The Second Incompleteness Theorem
- The Speedup Theorem
- The Continuum-Hypothesis Theorem
- The Time-Travel Theorem
- Gödel’s “God Theorem”
- Could a Finite Machine Match Gödel’s Greatness?






STOP & REVIEW IF NEEDED!

# Gödel's *Great Theorems* (OUP)

by Selmer Bringsjord


- Introduction (“The Wager”)
- Brief Preliminaries (e.g. the propositional calculus & FOL)
- The Completeness Theorem
- The First Incompleteness Theorem
-  The Second Incompleteness Theorem
- The Speedup Theorem
- The Continuum-Hypothesis Theorem
- The Time-Travel Theorem
- Gödel's “God Theorem”
- Could a Finite Machine Match Gödel's Greatness?



A corollary of the First Incompleteness Theorem: *We cannot prove (in “classical” mathematics) that mathematics is consistent.*

# Gödel's *Great Theorems* (OUP)

by Selmer Bringsjord

- Introduction (“The Wager”)
- Brief Preliminaries (e.g. the propositional calculus & FOL)
- The Completeness Theorem
- The First Incompleteness Theorem
-  The Second Incompleteness Theorem
- The Speedup Theorem
- The Continuum-Hypothesis Theorem
- The Time-Travel Theorem
- Gödel's “God Theorem”
- Could a Machine Match Gödel's Genius?



# Gödel's *Great Theorems* (OUP)

by Selmer Bringsjord

- Introduction (“The Wager”)
- Brief Preliminaries (e.g. the propositional calculus & FOL)
- The Completeness Theorem
- The First Incompleteness Theorem
- ➔ • The Second Incompleteness Theorem
- The Speedup Theorem
- ➔ • The Continuum-Hypothesis Theorem
- The Time-Travel Theorem
- Gödel’s “God Theorem”
- Could a Machine Match Gödel’s Genius?



By far the greatest of GGT; Selm’s analysis based Sherlock Holmes’ mystery “Silver Blaze.”

# The “Gödelian” Liar (from me)

# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

Suppose that  $\bar{P}$  is true. Then we can immediately deduce that  $\bar{P}$  is provable, because here is a proof:  $\bar{P} \rightarrow \bar{P}$  is an easy theorem, and from it and our supposition we deduce  $\bar{P}$  by *modus ponens*. But since what  $\bar{P}$  says is that it's unprovable, we have deduced that  $\bar{P}$  is false under our initial supposition.

# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

Suppose that  $\bar{P}$  is true. Then we can immediately deduce that  $\bar{P}$  is provable, because here is a proof:  $\bar{P} \rightarrow \bar{P}$  is an easy theorem, and from it and our supposition we deduce  $\bar{P}$  by *modus ponens*. But since what  $\bar{P}$  says is that it's unprovable, we have deduced that  $\bar{P}$  is false under our initial supposition.

Suppose on the other hand that  $\bar{P}$  is false. Then we can immediately deduce that  $\bar{P}$  is unprovable: Suppose for *reductio* that  $\bar{P}$  is provable; then  $\bar{P}$  holds as a result of some proof, but what  $\bar{P}$  says is that it's unprovable; and so we have contradiction. But since what  $\bar{P}$  says is that it's unprovable, and we have just proved that under our supposition, we arrive at the conclusion that  $\bar{P}$  is true.

# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

Suppose that  $\bar{P}$  is true. Then we can immediately deduce that  $\bar{P}$  is provable, because here is a proof:  $\bar{P} \rightarrow \bar{P}$  is an easy theorem, and from it and our supposition we deduce  $\bar{P}$  by *modus ponens*. But since what  $\bar{P}$  says is that it's unprovable, we have deduced that  $\bar{P}$  is false under our initial supposition.

Suppose on the other hand that  $\bar{P}$  is false. Then we can immediately deduce that  $\bar{P}$  is unprovable: Suppose for *reductio* that  $\bar{P}$  is provable; then  $\bar{P}$  holds as a result of some proof, but what  $\bar{P}$  says is that it's unprovable; and so we have contradiction. But since what  $\bar{P}$  says is that it's unprovable, and we have just proved that under our supposition, we arrive at the conclusion that  $\bar{P}$  is true.

$T(\bar{P})$  iff (i.e., if & only if)  $\neg T(\bar{P}) = F(\bar{P})$



# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

Suppose that  $\bar{P}$  is true. Then we can immediately deduce that  $\bar{P}$  is provable, because here is a proof:  $\bar{P} \rightarrow \bar{P}$  is an easy theorem, and from it and our supposition we deduce  $\bar{P}$  by *modus ponens*. But since what  $\bar{P}$  says is that it's unprovable, we have deduced that  $\bar{P}$  is false under our initial supposition.

Suppose on the other hand that  $\bar{P}$  is false. Then we can immediately deduce that  $\bar{P}$  is unprovable: Suppose for *reductio* that  $\bar{P}$  is provable; then  $\bar{P}$  holds as a result of some proof, but what  $\bar{P}$  says is that it's unprovable; and so we have contradiction. But since what  $\bar{P}$  says is that it's unprovable, and we have just proved that under our supposition, we arrive at the conclusion that  $\bar{P}$  is true.

$T(\bar{P})$  iff (i.e., if & only if)  $\neg T(\bar{P}) = F(\bar{P})$

Contradiction!

# The “Gödelian” Liar (from me)

$\bar{P}$ : This sentence is unprovable.

$\bar{\pi}$  = “I’m unprovable.”

Suppose that  $\bar{P}$  is true. Then we can immediately deduce that  $\bar{P}$  is provable, because here is a proof:  $\bar{P} \rightarrow \bar{P}$  is an easy theorem, and from it and our supposition we deduce  $\bar{P}$  by *modus ponens*. But since what  $\bar{P}$  says is that it’s unprovable, we have deduced that  $\bar{P}$  is false under our initial supposition.

Suppose on the other hand that  $\bar{P}$  is false. Then we can immediately deduce that  $\bar{P}$  is unprovable: Suppose for *reductio* that  $\bar{P}$  is provable; then  $\bar{P}$  holds as a result of some proof, but what  $\bar{P}$  says is that it’s unprovable; and so we have contradiction. But since what  $\bar{P}$  says is that it’s unprovable, and we have just proved that under our supposition, we arrive at the conclusion that  $\bar{P}$  is true.

$T(\bar{P})$  iff (i.e., if & only if)  $\neg T(\bar{P}) = F(\bar{P})$

Contradiction!



All of this is fishy; but Gödel, as we've seen, transformed it (by e.g. use of his encryption scheme) into utterly precise, impactful, indisputable reasoning ...

# **PA** (Peano Arithmetic):

$$\text{A1} \quad \forall x(0 \neq s(x))$$

$$\text{A2} \quad \forall x \forall y (s(x) = s(y) \rightarrow x = y)$$

$$\text{A3} \quad \forall x (x \neq 0 \rightarrow \exists y (x = s(y)))$$

$$\text{A4} \quad \forall x (x + 0 = x)$$

$$\text{A5} \quad \forall x \forall y (x + s(y) = s(x + y))$$

$$\text{A6} \quad \forall x (x \times 0 = 0)$$

$$\text{A7} \quad \forall x \forall y (x \times s(y) = (x \times y) + x)$$

And, every sentence that is the universal closure of an instance of

$$([\phi(0) \wedge \forall x(\phi(x) \rightarrow \phi(s(x)))] \rightarrow \forall x \phi(x))$$

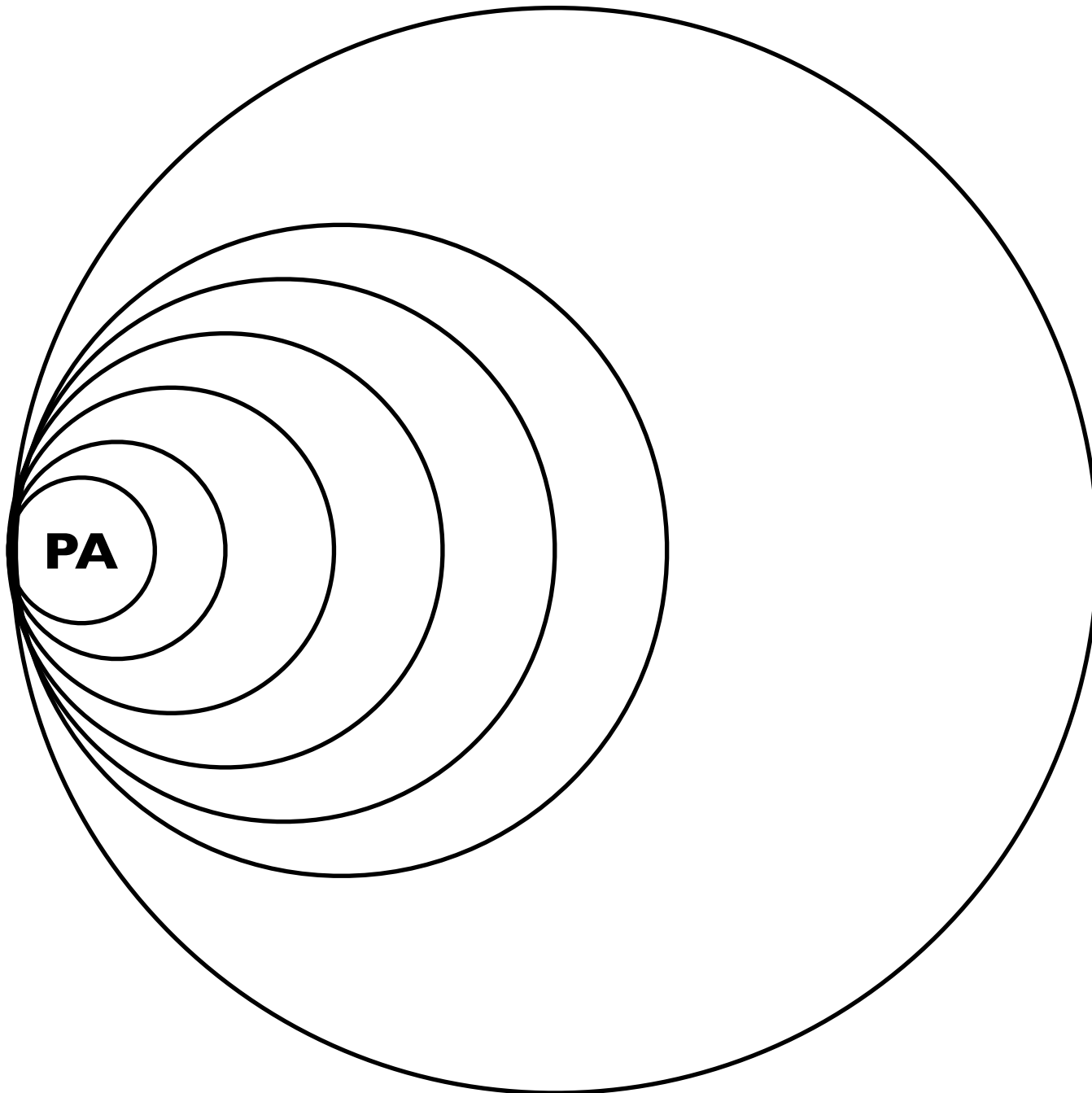
where  $\phi(x)$  is open wff with variable  $x$ , and perhaps others, free.

# Is there buried inconsistency in here?!?

Courtesy of Gödel: Given certain limitative assumptions about “proof power,”  
we can't prove that there isn't!

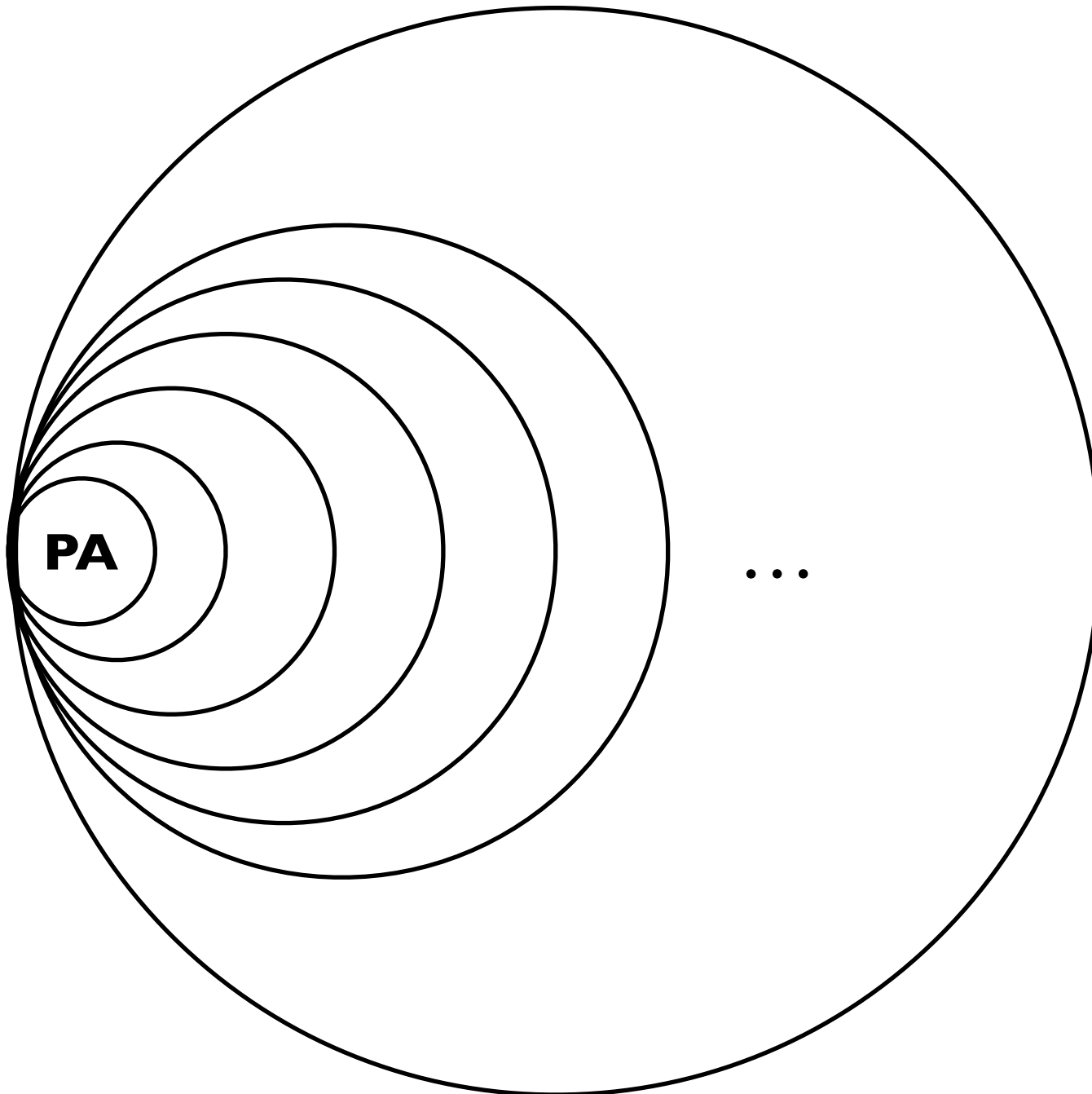
# Is there buried inconsistency in here?!?

Courtesy of Gödel: Given certain limitative assumptions about “proof power,”  
we can't prove that there isn't!



# Is there buried inconsistency in here?!?

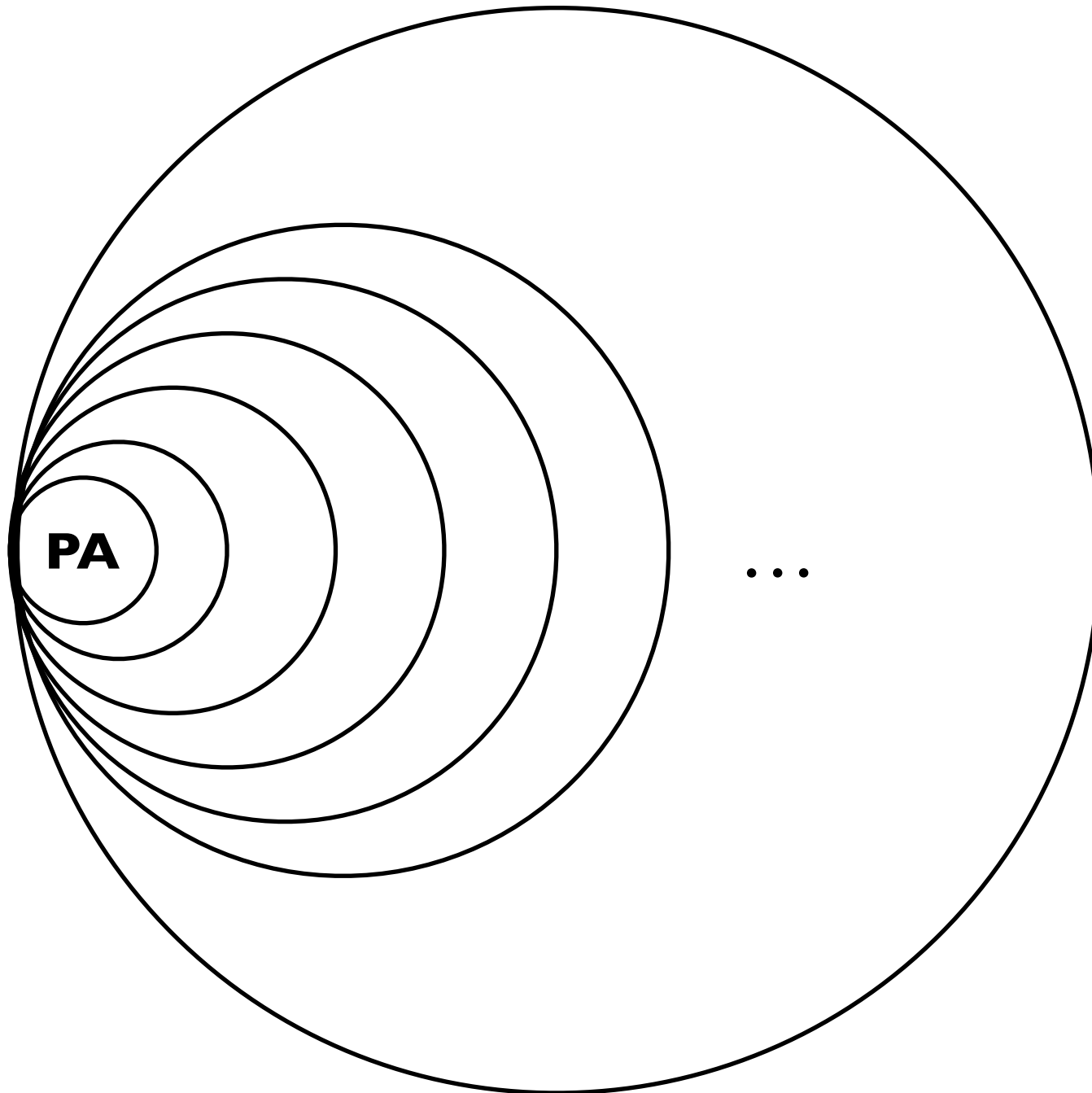
Courtesy of Gödel: Given certain limitative assumptions about “proof power,”  
we can't prove that there isn't!





# Is there buried inconsistency in here???

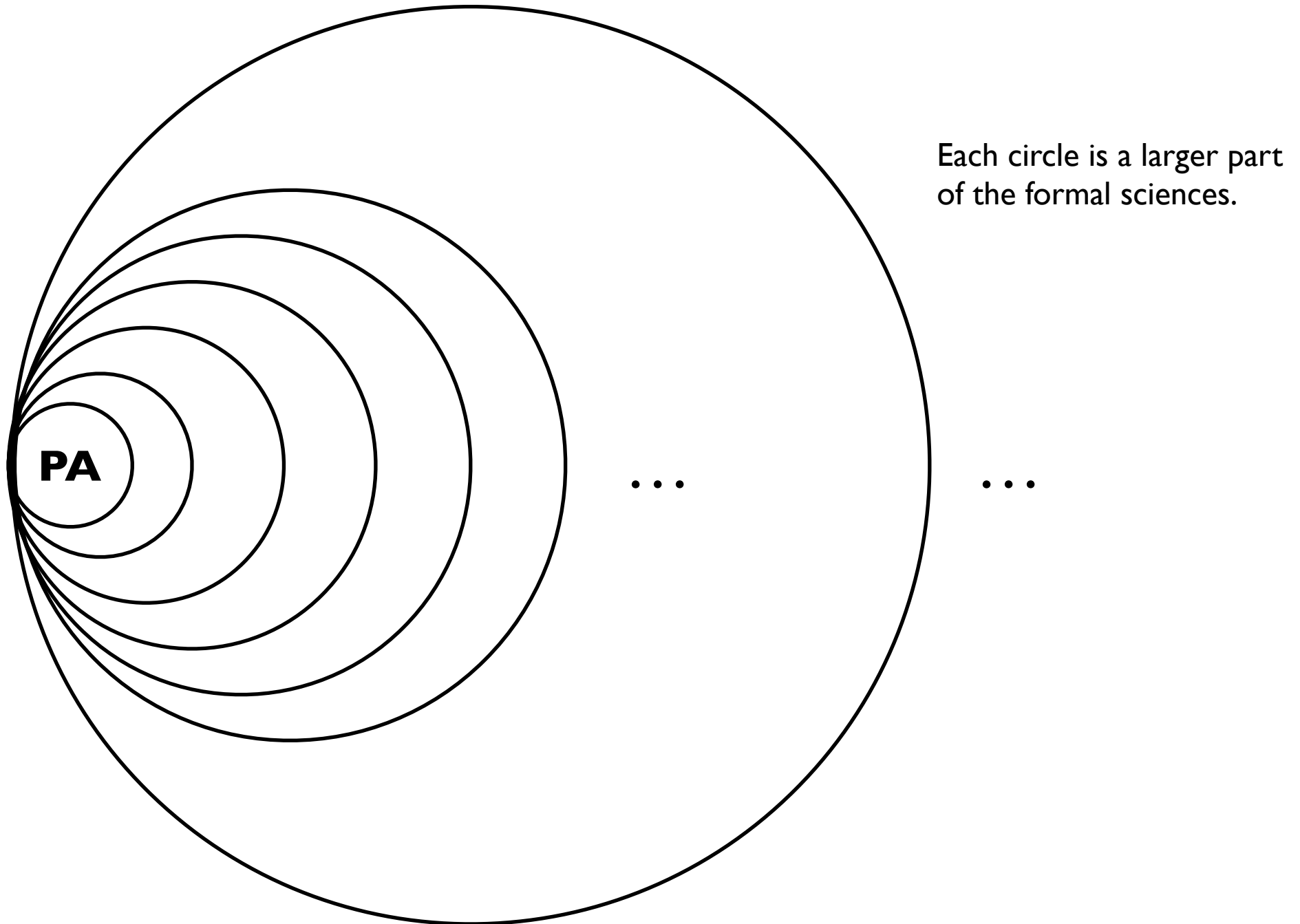
Courtesy of Gödel: Given certain limitative assumptions about “proof power,”  
we can't prove that there isn't!



Each circle is a larger part  
of the formal sciences.

# Is there buried inconsistency in here???

Courtesy of Gödel: Given certain limitative assumptions about “proof power,”  
we can't prove that there isn't!



**G2 as Slogan ...**

# **G2 as Slogan ...**

“We can't use math to ascertain whether mathematics is consistent.”

**G2 as Slogan ...**

# G2 as Slogan ...

“If we are restricted to certain kinds of formal reasoning, and feel we must have all of **PA** (math, engineering, etc.), we can't ascertain whether mathematics is consistent.”

# **Gödel's Second Incompleteness Theorem**

# Gödel's Second Incompleteness Theorem

Suppose  $\Phi \supset \mathbf{PA}$  that is

- (i) Con  $\Phi$ ;
- (ii) Turing-decidable (i.e. membership in  $\Phi$  is Turing-decidable); and
- (iii) sufficiently expressive to capture all of the operations of a Turing machine (i.e. Repr  $\Phi$ ).

Then  $\Phi \not\vdash \text{consis}_\Phi$ .



# Gödel's Second Incompleteness Theorem

Remember Church's Theorem!

Suppose  $\Phi \supset \mathbf{PA}$  that is

- (i) Con  $\Phi$ ;
- (ii) Turing-decidable (i.e. membership in  $\Phi$  is Turing-decidable); and
- (iii) sufficiently expressive to capture all of the operations of a Turing machine (i.e. Repr  $\Phi$ ).

Then  $\Phi \not\vdash \text{consis}_\Phi$ .

# Gödel's Second Incompleteness Theorem

Suppose  $\Phi \supset \mathbf{PA}$  that is

Remember Church's Theorem!

- (i) Con  $\Phi$ ;
- (ii) Turing-decidable (i.e. membership in  $\Phi$  is Turing-decidable); and
- (iii) sufficiently expressive to capture all of the operations of a Turing machine (i.e. Repr  $\Phi$ ).

Then  $\Phi \not\vdash \text{consis}_\Phi$ .

To prove  $G2$ , we shall once  
again allow ourselves ...

# The Fixed Point Theorem (FPT)

Assume that  $\Phi$  is a set of arithmetic sentences such that  $\text{Repr } \Phi$ . Then for every arithmetic formula  $\psi(x)$  with one free variable  $x$ , there is an arithmetic sentence  $\phi$  s.t.

$$\Phi \vdash \phi \leftrightarrow \psi(\hat{n}^\phi).$$

We can intuitively understand  $\phi$  to be saying:  
“I have the property ascribed to me by the formula  $\psi$ .”

# FPT in HyperSlate®!

The screenshot displays the HyperSlate interface with a dark top bar containing navigation icons and a status bar on the right that reads "FixedPointTheorem [HIGHER-ORDER-LOGIC]: Saved with 17 symbols." The main workspace shows a vertical flow of logical steps:

- A green box labeled "assume" is connected to a box containing the text "FIXED POINT THEOREM" in purple, followed by the formula  $\forall \psi: \exists \phi: \phi \Leftrightarrow \psi(\text{gn}(\phi))$  and the source "from {FIXED POINT THEOREM}" below it.
- An arrow points down to a green box labeled "v elim<sup>2</sup>".
- A final arrow points down to a box containing the text "I AM GREATER THAN ZERO" in purple, followed by the formula  $\exists \phi: \phi \Leftrightarrow \text{Greater}(\text{gn}(\phi), 0)$  and the source "from {FIXED POINT THEOREM}" below it.

Ok; so let's do it ... and let's see if you can see why Gödel declared  $G_2$  to be a direct “corollary” of  $G_1$ , and didn't bother to prove it in his original paper ...

**Proof:** Suppose that the antecedent (i)–(iii) of **G2** holds. Suppose for *reductio* that

$\Phi \vdash \text{consis}_{\Phi}$ .

We need three ingredients, and we shall be done. First, from FPT we can again directly obtain:

$$(*) \quad \Phi \vdash \mathcal{G} \leftrightarrow \neg \mathcal{P}_{\Phi}(\hat{n}^{\mathcal{G}}).$$

Next, we can prove (how? ... from one half of **G1!**) that:

$$(7.9) \quad \text{If } \text{Con } \Phi, \text{ then } \Phi \not\vdash \mathcal{G}.$$

Thirdly, we can logicize the meta-logical proposition that  $\Phi$  is consistent as an object-level conditional which can itself be proved formally from  $\Phi$ :

$$(**) \quad \Phi \vdash \text{consis}_{\Phi} \rightarrow \neg \mathcal{P}_{\Phi}(\hat{n}^{\mathcal{G}}).$$

Contradiction! (Can you find it?) **QED**





*Med nok penger, kan  
logikk løse alle problemer.*