Steeple #3:
Gödel's "Silver Blaze" Theorem

Selmer Bringsjord

Are Humans Rational?

Dec 8 2019

RPI
Troy NY USA

Rensselaer AI and Reasoning Lab

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- 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?

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- 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?

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- 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?

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- The Completeness Theorem

3/4

- 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?

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- 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?

Gödels' Great Theorems (OUP)

by Selmer Bringsjord

- Introduction ("The Wager")
- Brief Preliminaries (e.g. the propositional calculus)
- 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?

ZFC (The Foundation of Mathematics)

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

PAIR: Given two sets A and B, the pair set $\{A, B\}$ exists.

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

PAIR: Given two sets A and B, the pair set $\{A, B\}$ exists.

$\cdots$

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

PAIR: Given two sets A and B, the pair set $\{A, B\}$ exists.

$\ldots$

INF: There exists an infinite set, viz. $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \ldots\}$.

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

PAIR: Given two sets A and B, the pair set $\{A, B\}$ exists.

$\dots$

INF: There exists an infinite set, viz. $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \dots\}$.

AC: Given a set A of nonempty pairwise disjoint sets, there exists a set which contains exactly one element of each set in A.

ZFC (The Foundation of Mathematics)

EXT: Two sets which contain the same elements are equal.

PAIR: Given two sets A and B, the pair set $\{A, B\}$ exists.

$\cdots$

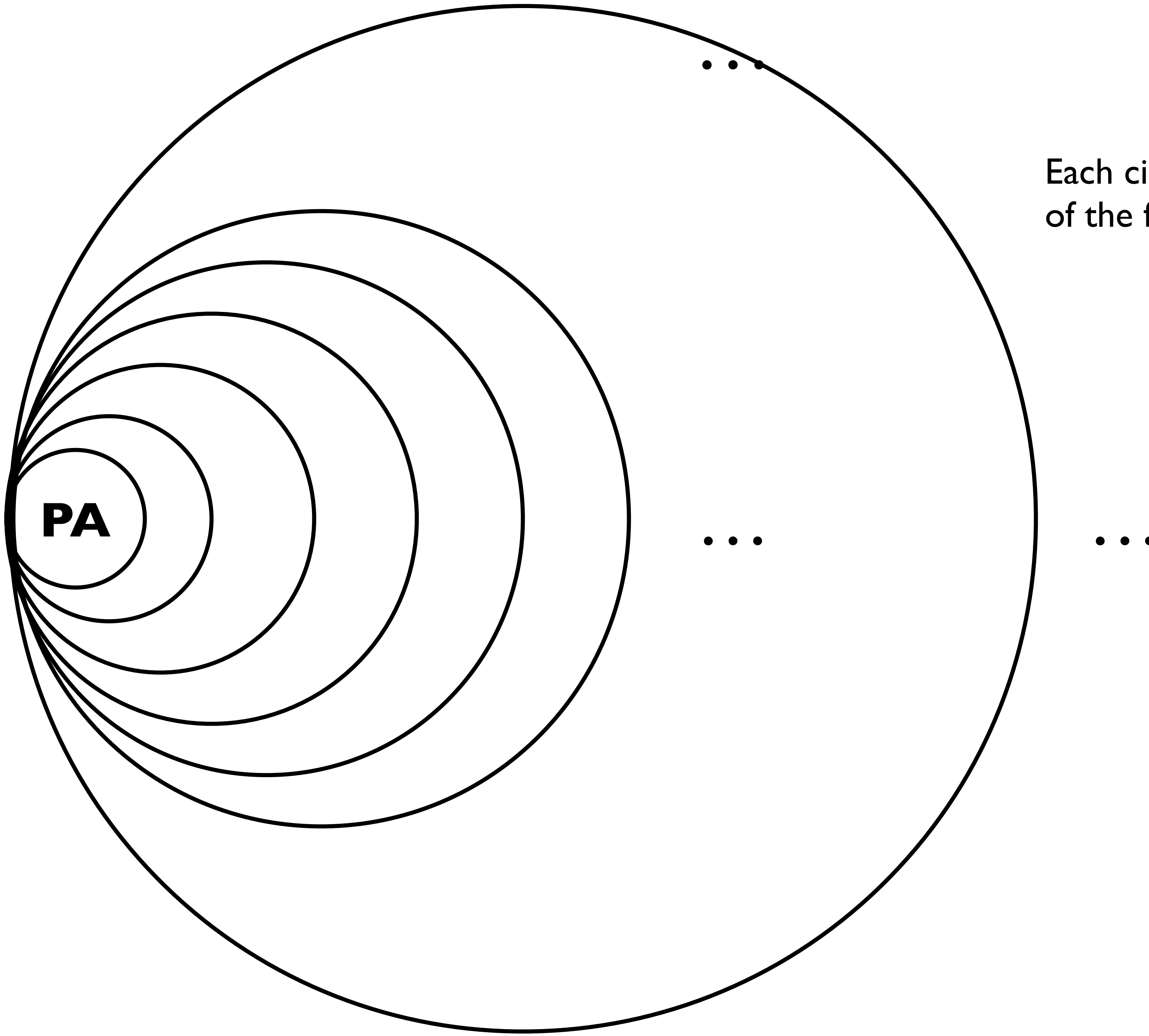
$\mathbb{N}$

INF: There exists an infinite set, viz. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots\}$.

AC: Given a set A of nonempty pairwise disjoint sets, there exists a set which contains exactly one element of each set in A.

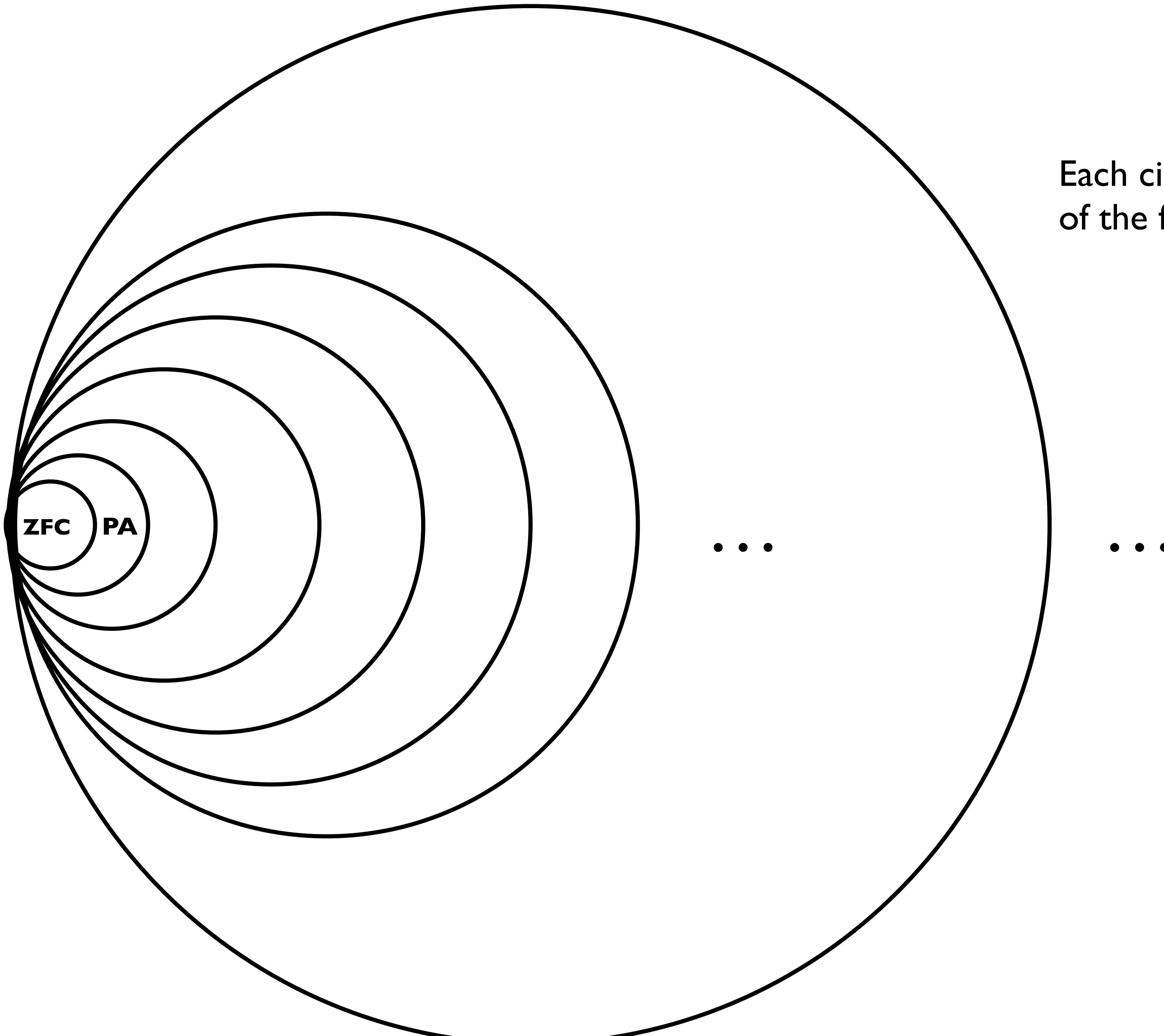
Arithmetic is Part of All Things Sci/Eng/Tech!

and courtesy of Gödel: We can't even prove all truths of arithmetic!



Each circle is a larger part of the formal sciences.

Actually, the true kernel is set theory!



Each circle is a larger part of the formal sciences.

Cantor’s Theorem

Cantor’s Theorem

Cantor (1878):

Cantor’s Theorem

Cantor (1878):

$$\mathcal{P}(\mathbb{N}) > \mathbb{N}$$

Cantor’s Theorem

Cantor (1878):

$$\mathcal{P}(\mathbb{N}) > \mathbb{N}$$

The power set of the natural numbers (i.e. the set of all subsets of the natural numbers) is larger than the natural numbers!

Cantor's Theorem

Cantor (1878):

$$\mathcal{P}(\mathbb{N}) > \mathbb{N}$$

The power set of the natural numbers (i.e. the set of all subsets of the natural numbers) is larger than the natural numbers!

How do we know this???!???

Continuum Hypothesis

Continuum Hypothesis

$$\mathrm{CH}: \forall S[(S \subset \mathbb{R} \wedge \neg \mathbf{Fin}(S)) \to (S \sim \mathbb{N} \vee S \sim \mathbb{R})]$$

Continuum Hypothesis

$$\mathrm{CH} : \forall S[(S \subset \mathbb{R} \land \neg \mathbf{Fin}(S)) \to (S \sim \mathbb{N} \lor S \sim \mathbb{R})]$$

Every infinite subset of the reals is either the same size as the natural numbers or the same size as the reals.

Generalized Continuum Hypothesis

Generalized Continuum Hypothesis

For every infinite set $S, \mathcal{P}(S) > S$.

Generalized Continuum Hypothesis

For every infinite set $S, \mathcal{P}(S) > S.$

Generalized Continuum Hypothesis (GCH):

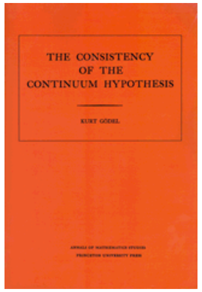
There's no set (size-wise) between *S* and *P*(*S*).

“Shorthand” History, Moral

Hilbert’s #1 (1900): “very plausible theorem”: **CH**

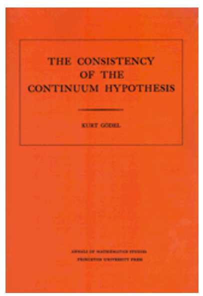
“Shorthand” History, Moral

Hilbert’s #1 (1900): “very plausible theorem”: **CH**



“Shorthand” History, Moral

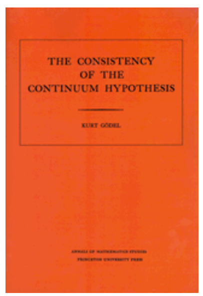
Hilbert’s #1 (1900): “very plausible theorem”: **CH**



Gödel (1938):

“Shorthand” History, Moral

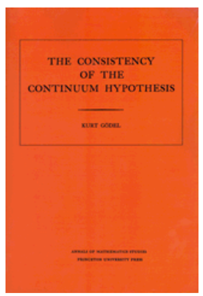
Hilbert’s #1 (1900): “very plausible theorem”: **CH**



Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

“Shorthand” History, Moral

Hilbert’s #1 (1900): “very plausible theorem”: **CH**

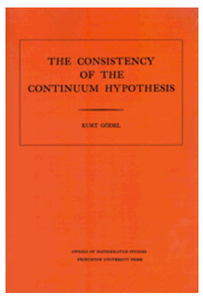


Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

Cohen (1963): $\mathbf{ZFC} \nvdash \mathbf{CH}$

“Shorthand” History, Moral

Hilbert’s #1 (1900): “very plausible theorem”: **CH**



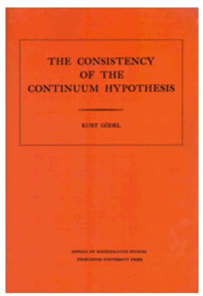
Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

Cohen (1963):
$\mathbf{ZFC} \nvdash \mathbf{CH}$

ADR-based approach to enabling an AI to discover and prove Gödel’s First Incompleteness Theorem:

“Shorthand” History, Moral

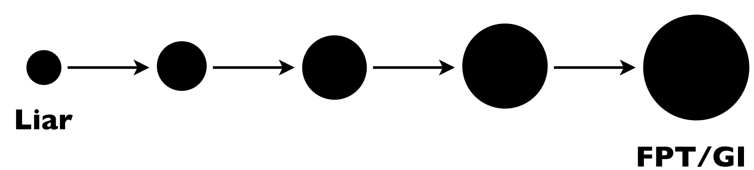
Hilbert’s #1 (1900): “very plausible theorem”: **CH**



Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

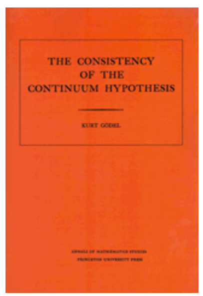
Cohen (1963):
$\mathbf{ZFC} \nvdash \mathbf{CH}$

ADR-based approach to enabling an AI to discover and prove Gödel’s First Incompleteness Theorem:



"Shorthand" History, Moral

Hilbert's #1 (1900): "very plausible theorem": **CH**



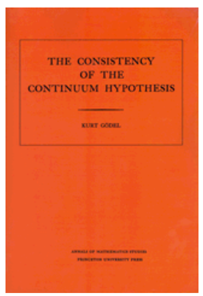
Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

Cohen (1963):
$\mathbf{ZFC} \nvdash \mathbf{CH}$

ADR-based approach to enabling an AI to discover and prove Gödel's First Incompleteness Theorem:

“Shorthand” History, Moral

Hilbert’s #1 (1900): “very plausible theorem”: **CH**



Gödel (1938): $\mathbf{ZFC} \nvdash \neg\mathbf{CH}$

Cohen (1963):
$\mathbf{ZFC} \nvdash \mathbf{CH}$

ADR-based approach to enabling an AI to discover and prove Gödel’s First Incompleteness Theorem:

Won’t work on this theorem of Gödel’s!

With Sherlock Holmes in the picture …

With Sherlock Holmes in the picture …

1900 Hilbert: “... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ...”

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.

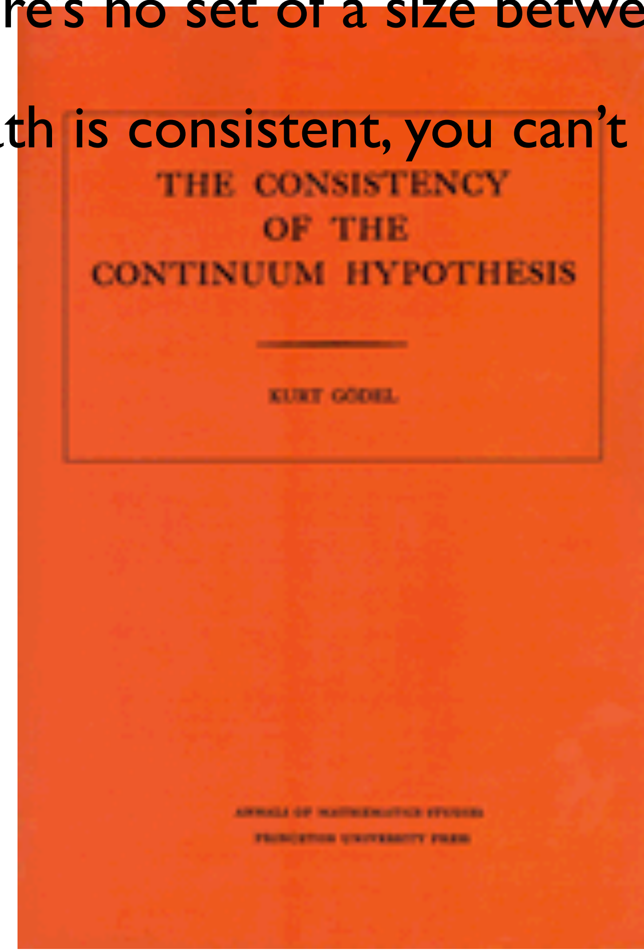
1938 Gödel: If math is consistent, you can't prove that CH is false.

With Sherlock Holmes in the picture …

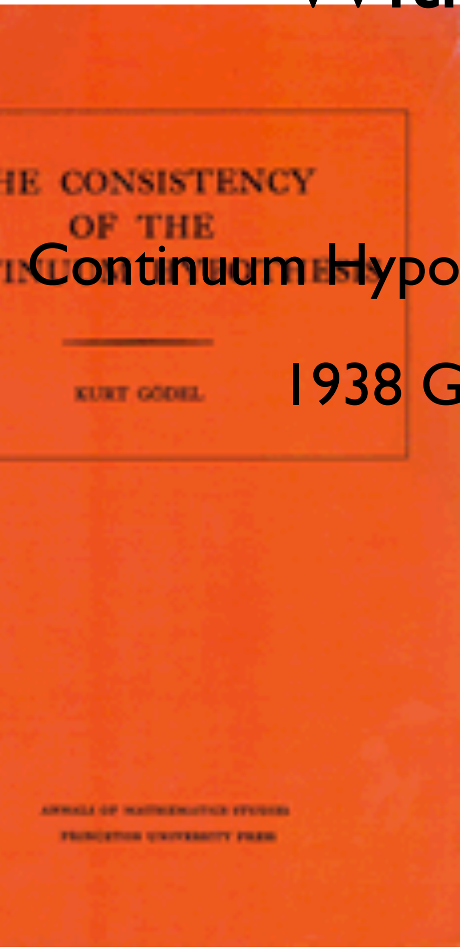
1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.

1938 Gödel: If math is consistent, you can't prove that CH is false.



With Sherlock Holmes in the picture …



t: "... a very plausible theorem, which nevertheless, in spite
strenuous efforts, no one has succeeded in proving. ..."

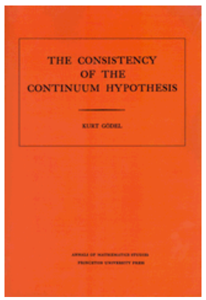
Continuum Hypothesis: There's no set of a size between the integers and the reals.

1938 Gödel: If math is consistent, you can't prove that CH is false.

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.

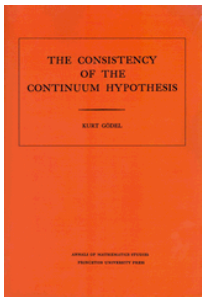


1938 Gödel: If math is consistent, you can't prove that CH is false.

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



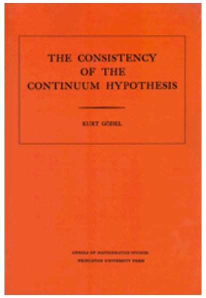
1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



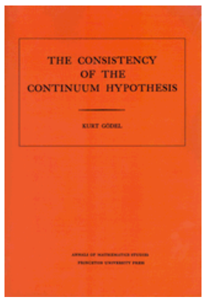
1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

With Sherlock Holmes in the picture …

1900 Hilbert: “... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ...”

Continuum Hypothesis: There’s no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can’t prove that CH is false.

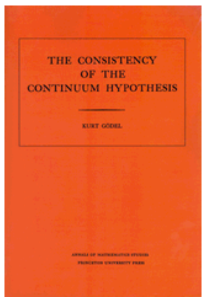
(1962 Cohen: If math is consistent, you can’t prove that CH is true.)

“Silver Blaze”

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can't prove that CH is false.

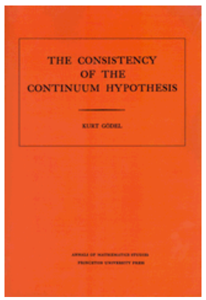
(1962 Cohen: If math is consistent, you can't prove that CH is true.)

"Silver Blaze"

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

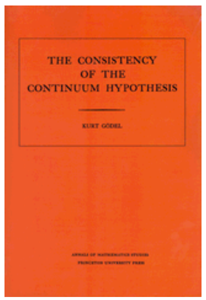
"Silver Blaze"

Straker dead from a "savage blow." Inspector Gregory:
"Fitzroy Simpson definitely guilty, and we arrested him!"

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

"Silver Blaze"

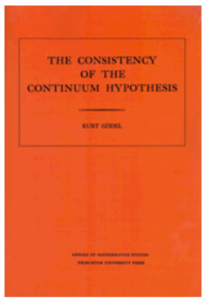
Straker dead from a "savage blow." Inspector Gregory: "Fitzroy Simpson definitely guilty, and we arrested him!"

Holmes: "But I've created a scenario in which all your clues/ evidence are true, but Fitzroy is perfectly innocent …"

With Sherlock Holmes in the picture …

1900 Hilbert: “... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ...”

Continuum Hypothesis: There’s no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can’t prove that CH is false.

(1962 Cohen: If math is consistent, you can’t prove that CH is true.)

“Silver Blaze”

Straker dead from a “savage blow.” Inspector Gregory: “Fitzroy Simpson definitely guilty, and we arrested him!”

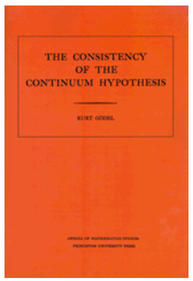
Holmes: “But I’ve created a scenario in which all your clues/ evidence are true, but Fitzroy is perfectly innocent …”

Gödel: “I’ve created a scenario where all of ZFC is true, but so is the Continuum Hypothesis!”

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

"Silver Blaze"

Straker dead from a "savage blow." Inspector Gregory: "Fitzroy Simpson definitely guilty, and we arrested him!"

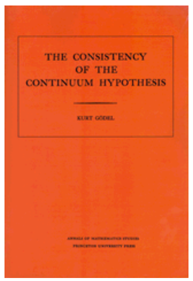
Holmes: "But I've created a scenario in which all your clues/ evidence are true, but Fitzroy is perfectly innocent …"

Gödel: "I've created a scenario where all of ZFC is true, but so is the Continuum Hypothesis!"

With Sherlock Holmes in the picture …

1900 Hilbert: "... a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. ..."

Continuum Hypothesis: There's no set of a size between the integers and the reals.



1938 Gödel: If math is consistent, you can't prove that CH is false.

(1962 Cohen: If math is consistent, you can't prove that CH is true.)

"Silver Blaze"

Straker dead from a "savage blow." Inspector Gregory: "Fitzroy Simpson definitely guilty, and we arrested him!"

Holmes: "But I've created a scenario in which all your clues/ evidence are true, but Fitzroy is perfectly innocent …"

Gödel: "I've created a scenario where all of ZFC is true, but so is the Continuum Hypothesis!"

Facts F …

Scenario G
(in honor of Inspector Gregory)

"And yet," said I, "even now I fail to understand what the theory of the police can be."

"I am afraid that whatever theory we state has very grave objections to it," returned my companion. "The police imagine, I take it, that this Fitzroy Simpson, having drugged the lad, and having in some way obtained a duplicate key, opened the stable door and took out the horse, with the intention, apparently, of kidnapping him altogether. His bridle is missing, so that Simpson must have put this on. Then, having left the door open behind him, he was leading the horse away over the moor, when he was either met or overtaken by the trainer. A row naturally ensued. Simpson beat out the trainer's brains with his heavy stick without receiving any injury from the small knife which Straker used in self-defense, and then the thief either led the horse on to some secret hiding-place, or else it may have bolted during the struggle, and be now wandering out on the moors. That is the case as it appears to the police, and improbable as it is, all other explanations are more improbable still. However, I shall very quickly test the matter when I am once upon the spot, and until then I cannot really see how we can get much further than our present position."

Sherlock as Logician

Sherlock as Logician

F = facts of the case
G = Inspector Gregory's scenario
H = Holmes's scenario

Sherlock as Logician

F = facts of the case
G = Inspector Gregory's scenario
H = Holmes's scenario

Holmes: "Gregory's claim is G (Simpson is guilty, & other details re. what he did). We have F, the facts of the case, disputed by no one. The question is: F $\vdash$ G? The answer is clearly No, for my scenario H, which entails $\neg G$, is consistent with the facts [Con $(F \cup H)$], and H entails $\neg G$. Here's the proof: Suppose for *reductio* that F $\vdash$ G. Then F $\cup$ H are inconsistent — contradiction! Hence F $\nvdash$ G."

Gödel's "Holmesian" Proof

$\mathrm{ZFC} \nvdash \mathrm{GCH}$

Gödel's "Holmesian" Proof

$\text{ZFC} \nvdash \text{GCH}$

Proof-Sketch: GCH is equivalent to the statement that there are at most *c** subsets of *c*, where '*c*' is any cardinal number, and *c** is the cardinal after *c* (i.e., $|P(c)| \leq c^*$). Gödel showed by employing transfinite recursion (!) that there are at most *c** constructible subsets of *c;* hence if all all sets are constructible, GCH follows. Gödel created a scenario in which all sets *are* constructible (viz., C), and showed that this scenario is consistent with ZFC (assuming, as we are, that ZFC is consistent). Therefore, by parallel to Holmes's reasoning, GCH can never be disproved from ZFC! Here''s the reasoning: Suppose for reductio that $\text{ZFC} \vdash \neg\text{GCH}$. We know that $\text{C} \vdash \text{GCH}$. Therefore C and ZFC are inconsistent; but this contradicts the consistency that Gödel proved. **QED**

Gödel's "Holmesian" Proof

ZFC $\nvdash$ GCH

Proof-Sketch: GCH is equivalent to the statement that there are at most *c** subsets of *c*, where '*c*' is any cardinal number, and *c** is the cardinal after *c* (i.e., $|P(c)| \leq c^*$). Gödel showed by employing transfinite recursion (!) that there are at most *c** constructible subsets of *c;* hence if all all sets are constructible, GCH follows. Gödel created a scenario in which all sets *are* constructible (viz., C), and showed that this scenario is consistent with ZFC (assuming, as we are, that ZFC is consistent). Therefore, by parallel to Holmes's reasoning, GCH can never be disproved from ZFC! Here''s the reasoning: Suppose for reductio that ZFC $\vdash \neg$GCH. We know that C $\vdash$GCH. Therefore C and ZFC are inconsistent; but this contradicts the consistency that Gödel proved. **QED**

(Of course, in "Silver Blaze," Sherlock did proceed to show that his scenario was true. The analogue to that doesn't hold in the Gödel story.)

slutten

Test 3 …

Question 1 (required) ...

$\mathcal{R}$ Humans, at least when neurobiologically normal and sane, are potentially presently (i.e., to abbreviate this double-adverb mouthful, *fundamentally*) rational, where such rationality is operationalized by certain explicit logico-mathematically based reasoning and decision-making successfully provided in response to real-world stimuli, especially such stimuli as are given in the form of focused, academic-style tests. However, mere animals are *not* fundamentally rational, since, *contra* Darwin,[1] their minds — unable as they are to e.g. reason in accord with abstract inference schemata like *modus tollens* — are provably qualitatively inferior to the human mind, as empirically confirmed by their poor performance on relevant focused, academic-style tests. As to whether computing machines/AIs/robots are fundamentally rational, the answer is also "No." For starters, if x can't read, really learn, write, and create, x can't be fundamentally rational; neither computing machines/robots nor non-human animals can read, really learn, write, or create; ergo, they aren't fundamentally rational for this reason alone. But the news for non-human animals and computing machines/robots gets much worse, for they have not the slightest chance when they are measured against the activity called out in $\mathcal{H}$.

$\mathcal{R}$ Humans, at least when neurobiologically normal and sane, are potentially presently (i.e., to abbreviate this double-adverb mouthful, *fundamentally*) rational, where such rationality is operationalized by certain explicit logico-mathematically based reasoning and decision-making successfully provided in response to real-world stimuli, especially such stimuli as are given in the form of focused, academic-style tests. However, mere animals are *not* fundamentally rational, since, *contra* Darwin,[1] their minds — unable as they are to e.g. reason in accord with abstract inference schemata like *modus tollens* — are provably qualitatively inferior to the human mind, as empirically confirmed by their poor performance on relevant focused, academic-style tests. As to whether computing machines/AIs/robots are fundamentally rational, the answer is also "No." For starters, if x can't read, really learn, write, and create, x can't be fundamentally rational; neither computing machines/robots nor non-human animals can read, really learn, write, or create; ergo, they aren't fundamentally rational for this reason alone. But the news for non-human animals and computing machines/robots gets much worse, for they have not the slightest chance when they are measured against the activity called out in $\mathcal{H}$.

1: No, humans aren't fundamentally rational as asserted here! & here's why!

2: No, nonhuman animals aren't non-rational as asserted here! & here's why!

3: No, AIs aren't non-rational as asserted here! & here's why!

4: No, the capacities called out in *H* are not beyond nonhuman animals and AIs.

Question 2 (optional) …

Indirect Proof

GI: g^* isn't provable from **PA**; nor is the negation of g^*!

Proof: Let's follow The Liar: Suppose that g^* *is* provable from PA; i.e., suppose **PA** |- g^*. Then by (1), with g* substituted for s, we have:

PA |- q^*("g^*") iff **PA** |- g^* (1')

From our supposition and working right to left by *modus ponens* on (1') we deduce:

PA |- q^*("g^*") (3.1)

But from our supposition and the earlier (see previous slide) (2), we can deduce by *modus ponens* that from PA the opposite can be proved! I.e., we have:

PA |- not-q^*("g^*") (3.2)

But (3.1) and (3.2) together means that **PA** is inconsistent, since it generates a contradiction. But we are working under the supposition that **PA** is consistent. Hence by indirect proof g^* is *not* provable from **PA**.

Indirect Proof

GI: g^* isn't provable from **PA**; nor is the negation of g^*!

Proof: Let's follow The Liar: Suppose that g^* *is* provable from PA; i.e., suppose **PA** |- g^*. Then by (1), with g* substituted for s, we have:

PA |- q^*("g^*") iff **PA** |- g^* (1')

From our supposition and working right to left by *modus ponens* on (1') we deduce:

PA |- q^*("g^*") (3.1)

But from our supposition and the earlier (see previous slide) (2), we can deduce by *modus ponens* that from PA the opposite can be proved! I.e., we have:

PA |- not-q^*("g^*") (3.2)

But (3.1) and (3.2) together means that **PA** is inconsistent, since it generates a contradiction. But we are working under the supposition that **PA** is consistent. Hence by indirect proof g^* is *not* provable from **PA**.

Show, in a manner that follows the reasoning in The Liar, that supposing that the negation of g^* (i.e., not-g^*) is provable from **PA** leads to a contradiction, and hence can't be provable from **PA**.

Paper Turn-In Logistics …

Email final draft of paper as a pdf (no other formats accepted) attachment to *both* Selmer.Bringsjord@gmail.com and can.mekik@gmail.com using verbatim the SUBJECT "AHR? S19 Paper Final Draft <firstname> <lastname>" by 1159pm Dec 17 2019.

Thank you. Its been an honor to teach this class.