On the hierarchy of natural theories

James Walsh

Sage School of Philosophy
Cornell University

jameswalsh@cornell.edu

4/6/22

The starting points are Gödel's incompleteness theorems.

The starting points are Gödel's incompleteness theorems.

Theorem (Gödel)

No reasonable axiomatic theory is complete.

The starting points are Gödel's incompleteness theorems.

Theorem (Gödel)

No reasonable axiomatic theory is complete.

Theorem (Gödel)

No reasonable axiomatic theory proves its own consistency.

The starting points are Gödel's incompleteness theorems.

Theorem (Gödel)

No reasonable axiomatic theory is complete.

Theorem (Gödel)

No reasonable axiomatic theory proves its own consistency.

No axiom system suffices for the development of all of mathematics; how should we navigate the vast array of axiomatic theories?

The so-called **consistency strength hierarchy** maps out the reasonable axiomatic theories and their relations.

The so-called **consistency strength hierarchy** maps out the reasonable axiomatic theories and their relations.

Definition

For a base theory B, we say that $T \leq^{B}_{\mathrm{Con}} U$ if B proves that the consistency of U implies the consistency of T.

The so-called **consistency strength hierarchy** maps out the reasonable axiomatic theories and their relations.

Definition

For a base theory B, we say that $T \leq^B_{\mathrm{Con}} U$ if B proves that the consistency of U implies the consistency of T.

Definition

$T <^B_{\mathrm{Con}} U$ if $T \leq^B_{\mathrm{Con}} U$ and $U \not\leq^B_{\mathrm{Con}} T$.

The so-called **consistency strength hierarchy** maps out the reasonable axiomatic theories and their relations.

Definition

For a base theory B, we say that $T \leq^B_{\mathrm{Con}} U$ if B proves that the consistency of U implies the consistency of T.

Definition

$T <^B_{\mathrm{Con}} U$ if $T \leq^B_{\mathrm{Con}} U$ and $U \not\leq^B_{\mathrm{Con}} T$.

Definition

T and U are **equiconsistent** over B if $T \leq^B_{\mathrm{Con}} U$ and $U \leq^B_{\mathrm{Con}} T$.

Theorem (Folklore)

*$<_{\mathrm{Con}}$ is not **pre-linear**, i.e., there are non-equiconsistent T and U such that $T \nless_{\mathrm{Con}} U$ and $U \nless_{\mathrm{Con}} T$.*

Theorem (Folklore)

$<_{\mathrm{Con}}$ *is not* ***pre-linear****, i.e., there are non-equiconsistent* T *and* U *such that* $T \not<_{\mathrm{Con}} U$ *and* $U \not<_{\mathrm{Con}} T$.

Theorem (Folklore)

The ordering $<_{\mathrm{Con}}$ *is* ***ill-founded****, i.e., there is a sequence* $T_0 >_{\mathrm{Con}} T_1 >_{\mathrm{Con}} T_2 >_{\mathrm{Con}} \ldots$ *where each* T_i *is consistent.*

All known instances of non-linearity and ill-foundedness are ad hoc; they were discovered by applying logical techniques.

All known instances of non-linearity and ill-foundedness are ad hoc; they were discovered by applying logical techniques.

Empirical Observation: The restriction of $<_{\mathrm{Con}}$ to the theories that arise in practice is a *well-ordering*.

$$\mathrm{EA}, \mathrm{EA}^{+}, \mathrm{PRA}, I\Sigma_n, \mathrm{PA}, \mathrm{ATR}_0, \Pi^1_n\mathrm{CA}_0, \mathrm{PA}_n, \mathrm{ZF}, \mathrm{AD}^{L(\mathbb{R})}$$

All known instances of non-linearity and ill-foundedness are ad hoc; they were discovered by applying logical techniques.

Empirical Observation: The restriction of $<_{\mathrm{Con}}$ to the theories that arise in practice is a *well-ordering*.

$$\mathrm{EA}, \mathrm{EA}^{+}, \mathrm{PRA}, I\Sigma_n, \mathrm{PA}, \mathrm{ATR}_0, \Pi^1_n\mathrm{CA}_0, \mathrm{PA}_n, \mathrm{ZF}, \mathrm{AD}^{L(\mathbb{R})}$$

Explaining this contrast is widely regarded as a major outstanding conceptual problem in mathematical logic.

The fact that "natural" theories, i.e. theories which have something like an "idea" to them, are almost always linearly ordered with regard to logical strength has been called one of the great mysteries of the foundations of mathematics.

S. Friedman, Rathjen, Weiermann

Three reasons for discussing set theory.

Three reasons for discussing set theory.

1. Set theory has proceeded in an explicitly axiomatic way since the isolation of ZFC.

Three reasons for discussing set theory.

1. Set theory has proceeded in an explicitly axiomatic way since the isolation of ZFC.
2. ZFC is highly general.

Three reasons for discussing set theory.

1. Set theory has proceeded in an explicitly axiomatic way since the isolation of ZFC.
2. ZFC is highly general.
3. ZFC is insufficient for answering many of the problems that motivated the early development of set theory:
    - The Continuum Hypothesis
    - Projective Measure
    - Suslin's Hypothesis

Set theorists have investigated a wide array of extensions of ZFC.

Set theorists have investigated a wide array of extensions of ZFC.

- large cardinal axioms

Set theorists have investigated a wide array of extensions of ZFC.

- large cardinal axioms
- determinacy axioms

Set theorists have investigated a wide array of extensions of ZFC.

- large cardinal axioms
- determinacy axioms
- forcing axioms

Set theorists have investigated a wide array of extensions of ZFC.

- large cardinal axioms
- determinacy axioms
- forcing axioms

Can we make rational judgments about the correctness of these principles or their consequences?

Set theorists have investigated a wide array of extensions of ZFC.

- large cardinal axioms
- determinacy axioms
- forcing axioms

Can we make rational judgments about the correctness of these principles or their consequences?

Steel has promoted the following maxim:

MAXIMIZE STRENGTH.

Steel's Maxim echoes Cantor's dictum of mathematical freedom.

Steel's Maxim echoes Cantor's dictum of mathematical freedom.

The $<_{\mathrm{Con}}$ tells us what mathematics can be developed on the basis of one theory rather than another; (more or less) if $\mathrm{Con}(T)$ implies $\mathrm{Con}(U)$ then T can **interpret** U and **not** vice-versa.

Steel's Maxim echoes Cantor's dictum of mathematical freedom.

The $<_{\text{Con}}$ tells us what mathematics can be developed on the basis of one theory rather than another; (more or less) if $\text{Con}(T)$ implies $\text{Con}(U)$ then T can **interpret** U and **not** vice-versa.

- Poincaré interpreted two dimensional hyperbolic geometry in the Euclidean geometry of the unit circle.

Steel's Maxim echoes Cantor's dictum of mathematical freedom.

The $<_{\mathrm{Con}}$ tells us what mathematics can be developed on the basis of one theory rather than another; (more or less) if $\mathrm{Con}(T)$ implies $\mathrm{Con}(U)$ then T can **interpret** U and **not** vice-versa.

- Poincaré interpreted two dimensional hyperbolic geometry in the Euclidean geometry of the unit circle.
- Dedekind interpreted analysis in set theory.

Steel's Maxim echoes Cantor's dictum of mathematical freedom.

The $<_{\mathrm{Con}}$ tells us what mathematics can be developed on the basis of one theory rather than another; (more or less) if $\mathrm{Con}(T)$ implies $\mathrm{Con}(U)$ then T can **interpret** U and **not** vice-versa.

- Poincaré interpreted two dimensional hyperbolic geometry in the Euclidean geometry of the unit circle.
- Dedekind interpreted analysis in set theory.
- Gödel interpreted proof theory in arithmetic.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

1. φ increases strength but $\neg\varphi$ does not.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

1. φ increases strength but $\neg\varphi$ does not.
2. $\neg\varphi$ increases strength but φ does not.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

1. φ increases strength but $\neg\varphi$ does not.
2. $\neg\varphi$ increases strength but φ does not.
3. Neither φ nor $\neg\varphi$ increases strength.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

1. φ increases strength but $\neg\varphi$ does not.
2. $\neg\varphi$ increases strength but φ does not.
3. Neither φ nor $\neg\varphi$ increases strength.
4. Both φ and $\neg\varphi$ increases strength.

Is Steel's Maxim coherent?

Let's consider some sentence φ that is independent of ZFC.

1. φ increases strength but $\neg\varphi$ does not.
2. $\neg\varphi$ increases strength but φ does not.
3. Neither φ nor $\neg\varphi$ increases strength.
4. Both φ and $\neg\varphi$ increases strength.

It turns out that **all four** possibilities are realized; in the fourth case we **cannot** follow Steel's Maxim.

1. φ increases strength but $\neg\varphi$ does not.
2. $\neg\varphi$ increases strength but φ does not.
3. neither φ nor $\neg\varphi$ increases strength.
4. both φ and $\neg\varphi$ increases strength.

When we restrict our attention to natural theories, only the first three possibilities are realized.

This is just to say that natural theories are linearly ordered by consistency strength.

Consider again the axiom systems extending ZFC:

- large cardinal axioms
- axioms of definable determinacy
- forcing axioms

These systems have different motivations, but they are well-ordered by consistency strength.

Consider again the axiom systems extending ZFC:

- large cardinal axioms
- axioms of definable determinacy
- forcing axioms

These systems have different motivations, but they are well-ordered by consistency strength.

They converge on statements about $\mathbb{N}$; in fact, they converge on statements about $\mathbb{R}$.

Consider again the axiom systems extending ZFC:

- large cardinal axioms
- axioms of definable determinacy
- forcing axioms

These systems have different motivations, but they are well-ordered by consistency strength.

They converge on statements about $\mathbb{N}$; in fact, they converge on statements about $\mathbb{R}$.

> *At the level of sentences about $\mathbb{R}$, we know of only one road upward. We are led to it many different ways.*
>
> Steel

Fix a sufficiently strong, sound, effectively axiomatized theory T, e.g., *elementary arithmetic*, *Peano Arithmetic*, ...

Fix a sufficiently strong, sound, effectively axiomatized theory T, e.g., *elementary arithmetic*, *Peano Arithmetic*, ...

T is incomplete by Gödel's first theorem; T does not prove Con_T by Gödel's second theorem.

Fix a sufficiently strong, sound, effectively axiomatized theory T, e.g., *elementary arithmetic*, *Peano Arithmetic*, ...

T is incomplete by Gödel's first theorem; T does not prove Con_T by Gödel's second theorem.

Are there any proper extensions of T that are strictly weaker than $T + \mathsf{Con}_T$?

Rosser introduced a trick whereby we can find sentences strictly weaker than Con_T.

Rosser introduced a trick whereby we can find sentences strictly weaker than Con_T.

$$T \vdash \Big(R_T \leftrightarrow \forall x \big(\mathrm{Pf}_T(x, \ulcorner R_T \urcorner) \rightarrow \exists y < x \mathrm{Pf}_T(y, \ulcorner \neg R_T \urcorner) \big) \Big)$$

R_T "says": If there are any proofs of R_T, then they are preceded by proofs of $\neg R_T$.

Rosser introduced a trick whereby we can find sentences strictly weaker than Con_T.

$$T \vdash \Big(R_T \leftrightarrow \forall x \big(\mathrm{Pf}_T(x, \ulcorner R_T \urcorner) \rightarrow \exists y < x \mathrm{Pf}_T(y, \ulcorner \neg R_T \urcorner) \big) \Big)$$

R_T "says": If there are any proofs of R_T, then they are preceded by proofs of $\neg R_T$.

We can use Rosser's trick to produce independent sentences strictly weaker than Con_T.

$$\mathrm{Con}_T \vee R_{T+\neg\mathrm{Con}_T}$$

Yet these sentences are highly unnatural.

Rosser introduced a trick whereby we can find sentences strictly weaker than Con_T.

$$T \vdash \Big(R_T \leftrightarrow \forall x \big(\mathrm{Pf}_T(x, \ulcorner R_T \urcorner) \rightarrow \exists y < x \mathrm{Pf}_T(y, \ulcorner \neg R_T \urcorner) \big) \Big)$$

R_T "says": If there are any proofs of R_T, then they are preceded by proofs of $\neg R_T$.

We can use Rosser's trick to produce independent sentences strictly weaker than Con_T.

$$\mathrm{Con}_T \vee R_{T+\neg \mathrm{Con}_T}$$

Yet these sentences are highly unnatural.

Self-reference, dependence on a seemingly arbitrary numeration of proofs,...

Instead of focusing on specific theories, we focus on algorithms for **uniformly** extending theories.

Instead of focusing on specific theories, we focus on algorithms for **uniformly** extending theories.

We are particularly interested in $\mathfrak{g}$ that are *monotone*:

$$\text{If } T \text{ proves } \varphi \rightarrow \psi, \text{ then } T \text{ proves } \mathfrak{g}(\varphi) \rightarrow \mathfrak{g}(\psi).$$

There are many monotone algorithms for uniformly extending theories.

There are many monotone algorithms for uniformly extending theories.

$\mathfrak{id} : \varphi \mapsto \varphi$

There are many monotone algorithms for uniformly extending theories.

$\mathfrak{id} : \varphi \mapsto \varphi$

$\mathfrak{Con} : \varphi \mapsto \mathrm{Con}_T(\varphi)$

There are many monotone algorithms for uniformly extending theories.

$\mathfrak{id} : \varphi \mapsto \varphi$

$\mathfrak{Con} : \varphi \mapsto \mathsf{Con}_T(\varphi)$

Rosser's trick engenders an algorithm for extending theories, but it is **not** monotone.

There are many monotone algorithms for uniformly extending theories.

$\mathfrak{id} : \varphi \mapsto \varphi$

$\mathfrak{Con} : \varphi \mapsto \mathsf{Con}_T(\varphi)$

Rosser's trick engenders an algorithm for extending theories, but it is **not** monotone.

Indeed, the Rosser algorithm is not monotone in virtue of the pathological properties flagged earlier.

The goal is to prove that the consistency operator is the unique weakest monotone algorithm for uniformly extending theories.

The goal is to prove that the consistency operator is the unique weakest monotone algorithm for uniformly extending theories.

What does "the unique weakest" mean?

We can make sense of this claim **only** modulo a suitable equivalence relation.

The goal is to prove that the consistency operator is the unique weakest monotone algorithm for uniformly extending theories.

What does "the unique weakest" mean?

We can make sense of this claim **only** modulo a suitable equivalence relation.

Let φ be a true sentence. Then the set of sentences that implies φ is a **cone**.

$$\{\psi : T + \psi \text{ proves } \varphi\}$$

Let's call a function $\mathfrak{g}$ bounded if there exists a $k \in \mathbb{N}$ such that, for every φ, $\mathfrak{g}(\varphi) \in \Pi^0_k$.

For technical reasons, we restrict our attention to bounded functions.

Theorem (W.)

Let $\mathfrak{g}$ be a bounded, computable, and monotone. Then one of the following holds:

1. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$, $T + \varphi \vdash \mathfrak{g}(\varphi)$.*
2. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$, $T + \varphi + \mathfrak{g}(\varphi) \vdash \mathrm{Con}_T(\varphi)$.*

Theorem (W.)

Let $\mathfrak{g}$ be a bounded, computable, and monotone. Then one of the following holds:

1. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$, $T + \varphi \vdash \mathfrak{g}(\varphi)$.*
2. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$,* $T + \varphi + \mathfrak{g}(\varphi) \vdash \mathrm{Con}_T(\varphi)$.

That is, either $\mathfrak{g}$ is as weak as the identity on a cone or as strong as the consistency operator on a cone.

Theorem (W.)

Let $\mathfrak{g}$ be a bounded, computable, and monotone. Then one of the following holds:

1. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$, $T + \varphi \vdash \mathfrak{g}(\varphi)$.*
2. *There is a cone $\mathfrak{C}$ such that for all $\varphi \in \mathfrak{C}$,* $T + \varphi + \mathfrak{g}(\varphi) \vdash \mathrm{Con}_T(\varphi)$.

That is, either $\mathfrak{g}$ is as weak as the identity on a cone or as strong as the consistency operator on a cone.

The consistency operator is the unique weakest method for uniformly extending theories.

This contributes to a partial explanation of the well-ordering phenomenon.

This contributes to a partial explanation of the well-ordering phenomenon.

It suggests that the iterates of the consistency operator form a spine of axiomatic theories that is, in some sense, canonical.

We now shift our attention to second-order arithmetic, the joint theory of the natural numbers and the real numbers.

We now shift our attention to second-order arithmetic, the joint theory of the natural numbers and the real numbers.

ACA_0 is our base theory; it is a second-order pendant of PA.

The Π^0_1 formulas are the formulas $\forall x \in \mathbb{N}\ \varphi$ where φ is computable.

The Π^0_1 formulas are the formulas $\forall x \in \mathbb{N}\ \varphi$ where φ is computable.

The Σ^0_1 formulas are the formulas $\exists x \in \mathbb{N}\ \varphi$ where φ is computable.

The Π^0_1 formulas are the formulas $\forall x \in \mathbb{N}\ \varphi$ where φ is computable.

The Σ^0_1 formulas are the formulas $\exists x \in \mathbb{N}\ \varphi$ where φ is computable.

The Π^1_1 formulas are the formulas $\forall x \in \mathbb{R}\ \varphi$ where φ has no quantifiers over $\mathbb{R}$.

The Π^0_1 formulas are the formulas $\forall x \in \mathbb{N}\ \varphi$ where φ is computable.

The Σ^0_1 formulas are the formulas $\exists x \in \mathbb{N}\ \varphi$ where φ is computable.

The Π^1_1 formulas are the formulas $\forall x \in \mathbb{R}\ \varphi$ where φ has no quantifiers over $\mathbb{R}$.

The Σ^1_1 formulas are the formulas $\exists x \in \mathbb{R}\ \varphi$ where φ has no quantifiers over $\mathbb{R}$.

Definition

A theory T is Γ-sound if every Γ sentence that T proves is true.

Definition

A theory T is Γ-sound if every Γ sentence that T proves is true.

$$\mathsf{RFN}_\Gamma(T) := \forall\varphi \in \Gamma\big(\mathsf{Pr}_T(\varphi) \to \mathsf{True}_\Gamma(\varphi)\big)$$

Definition

A theory T is Γ-sound if every Γ sentence that T proves is true.

$$\mathrm{RFN}_\Gamma(T) := \forall\varphi \in \Gamma\big(\mathrm{Pr}_T(\varphi) \rightarrow \mathrm{True}_\Gamma(\varphi)\big)$$

Fact: A theory is consistent just in case it is Π^0_1-sound.

Definition

A theory T is Γ-sound if every Γ sentence that T proves is true.

$$\mathrm{RFN}_\Gamma(T) := \forall \varphi \in \Gamma \left(\mathrm{Pr}_T(\varphi) \rightarrow \mathrm{True}_\Gamma(\varphi) \right)$$

Fact: A theory is consistent just in case it is Π^0_1-sound.

Definition

$T \vdash^\Gamma \varphi$ if there is a true $\psi \in \Gamma$ such that $T + \psi \vdash \varphi$.

Definition

A theory T is Γ-sound if every Γ sentence that T proves is true.

$$\mathsf{RFN}_\Gamma(T) := \forall\varphi \in \Gamma\big(\mathsf{Pr}_T(\varphi) \rightarrow \mathsf{True}_\Gamma(\varphi)\big)$$

Fact: A theory is consistent just in case it is Π^0_1-sound.

Definition

$T \vdash^\Gamma \varphi$ if there is a true $\psi \in \Gamma$ such that $T + \psi \vdash \varphi$.

Fact: For any T and φ, $T \vdash \varphi$ if and only if $T \vdash^{\Sigma^0_1} \varphi$.

Definition

$T \leq_{\mathrm{Con}} U := \mathrm{ACA}_0 \vdash \mathrm{Con}(U) \rightarrow \mathrm{Con}(T)$.

Definition

$T \leq_{\mathrm{Con}} U := \mathrm{ACA}_0 \vdash^{\Sigma^0_1} \mathrm{Con}(U) \rightarrow \mathrm{Con}(T).$

Definition

$$T \leq_{\mathrm{Con}} U := \mathrm{ACA}_0 \vdash^{\Sigma^0_1} \mathrm{RFN}_{\Pi^0_1}(U) \rightarrow \mathrm{RFN}_{\Pi^0_1}(T).$$

Definition

$$T \leq_{\mathsf{Con}} U := \mathsf{ACA}_0 \vdash^{\Sigma^0_1} \mathsf{RFN}_{\Pi^0_1}(U) \rightarrow \mathsf{RFN}_{\Pi^0_1}(T).$$

Definition

$$T \leq^{\Sigma^1_1}_{\Pi^1_1} U := \mathsf{ACA}_0 \vdash^{\Sigma^1_1} \mathsf{RFN}_{\Pi^1_1}(U) \rightarrow \mathsf{RFN}_{\Pi^1_1}(T).$$

Definition

$T \leq_{\mathrm{Con}} U := \mathrm{ACA}_0 \vdash^{\Sigma^0_1} \mathrm{RFN}_{\Pi^0_1}(U) \rightarrow \mathrm{RFN}_{\Pi^0_1}(T).$

Definition

$T \leq^{\Sigma^1_1}_{\Pi^1_1} U := \mathrm{ACA}_0 \vdash^{\Sigma^1_1} \mathrm{RFN}_{\Pi^1_1}(U) \rightarrow \mathrm{RFN}_{\Pi^1_1}(T).$

Theorem (W.)

The relation $\leq^{\Sigma^1_1}_{\Pi^1_1}$ *pre-well-orders the* Π^1_1*-sound extensions of* ACA_0.

Thanks!

J. Walsh (2020)
A note on the consistency operator.
Proceedings of the American Mathematical Society. 148(6):2645–2654

J. Walsh (2022)
On the hierarchy of natural theories.
arXiv.

J. Walsh (2022)
A robust proof-theoretic well-ordering.
arXiv.