This Title is False

Gödel's First Incompleteness Theorem

Hayden Jananthan

Vanderbilt University

Theorem (Gödel's First Incompleteness Theorem)

There are true sentences of arithmetic which are not provable.

Theorem (Gödel's First Incompleteness Theorem)

There are true sentences of arithmetic which are not provable.

1. What do we mean by 'sentences of arithmetic'?

Theorem (Gödel's First Incompleteness Theorem)

There are true sentences of arithmetic which are not provable.

1. What do we mean by 'sentences of arithmetic'?
2. What do we mean by 'true'?

Theorem (Gödel's First Incompleteness Theorem)

There are true sentences of arithmetic which are not provable.

1. What do we mean by 'sentences of arithmetic'?
2. What do we mean by 'true'?
3. What do we mean by 'provable'?

1. Meaning of Gödel's First Incompleteness Theorem

2. Proving Gödel's First Incompleteness Theorem

3. Generalizing Gödel's First Incompleteness Theorem

The Language of Arithmetic - Intuition

What are ‘sentences of arithmetic’?

The Language of Arithmetic - Intuition

What are ‘sentences of arithmetic’?

They are certain *well-formed statements in the language of arithmetic*, e.g.

The Language of Arithmetic - Intuition

What are 'sentences of arithmetic'?

They are certain *well-formed statements in the language of arithmetic*, e.g.

- "Addition is associative."
- "Every natural number greater than one is divisible by a prime number."
- "0 is not equal to $n+1$ for any natural number n."
- "For every two natural numbers n, m, a greatest common divisor $\gcd(n, m)$ exists."

The Language of Arithmetic - Formal Description

Formally, we define *sentences* as certain strings of symbols.
There are two kinds of symbols:

The Language of Arithmetic - Formal Description

Formally, we define *sentences* as certain strings of symbols.
There are two kinds of symbols:

Logical Symbols: Theory-independent symbols.

$\wedge$	(and),	$\vee$	(or),
$\rightarrow$	(implies),	$\leftrightarrow$	(if and only if),
$\neg$	(not),		
$\bot$	(falsehood),	$\top$	(truth),
$\forall$	(for all),	$\exists$	(there exists),
$\approx$	(equals),		
(	(left parentheses),	)	(right parentheses)
$x_0, x_1, x_2, \ldots$	(variables)		

Non-Logical Symbols: Theory-dependent symbols.

0	(zero),	S	(successor),
$+$,	(addition),	$\cdot$	(multiplication)

Terms and Well-Formed Formulas

Terms are defined recursively:

1. 0 and variables are terms.
2. If t_1, t_2 are terms, the following are terms:

$$S(t_1), \quad (t_1 + t_2), \quad (t_1 \cdot t_2)$$

Terms have *unique readability*: a term can be built up from earlier terms in exactly one way.

Terms and Well-Formed Formulas

Terms are defined recursively:

1. 0 and variables are terms.
2. If t_1, t_2 are terms, the following are terms:

$$S(t_1), \quad (t_1 + t_2), \quad (t_1 \cdot t_2)$$

Terms have *unique readability*: a term can be built up from earlier terms in exactly one way.

Well-formed formulas (wff) are defined recursively:

1. If t_1, t_2 are terms, then $(t_1 \approx t_2)$ is a wff.
2. $\top$ and $\bot$ are wffs.
3. If φ_1, φ_2 are wffs and x a variable, the following are wffs:

$$(\varphi_1 \wedge \varphi_2), \quad (\varphi_1 \vee \varphi_2), \quad (\varphi_1 \rightarrow \varphi_2), \quad (\varphi_1 \leftrightarrow \varphi_2), \quad \neg\varphi_1, \quad \forall x \varphi_1, \quad \exists x \varphi_1$$

Wffs have *unique readability*: a wff can be built up from earlier wffs and terms in exactly one way.

Sentences

Definition

An *instance* of a variable x in a wff φ is **bound** if it is contained in a substring of φ of the form

$$\forall x\psi \quad \text{or} \quad \exists x\psi$$

and **free** otherwise.

Sentences

Definition

An *instance* of a variable x in a wff φ is **bound** if it is contained in a substring of φ of the form

$$\forall x\psi \quad \text{or} \quad \exists x\psi$$

and **free** otherwise.

Definition

A **sentence** is a wff with no free variables.

Sentences

Definition

An *instance* of a variable x in a wff φ is **bound** if it is contained in a substring of φ of the form

$$\forall x \psi \quad \text{or} \quad \exists x \psi$$

and **free** otherwise.

Definition

A **sentence** is a wff with no free variables.

E.g.

- *Yes*: $\forall x \forall y \forall z (((x + y) + z) \approx (x + (y + z)))$ (associativity of $+$)
- *No*: $\exists x (x \cdot x = y)$ (y is a perfect square)

Sentences

Definition

An *instance* of a variable x in a wff φ is **bound** if it is contained in a substring of φ of the form

$$\forall x\psi \quad \text{or} \quad \exists x\psi$$

and **free** otherwise.

Definition

A **sentence** is a wff with no free variables.

E.g.

- *Yes*: $\forall x\forall y\forall z(((x+y)+z)\approx(x+(y+z)))$ (associativity of $+$)
- *No*: $\exists x(x\cdot x=y)$ (y is a perfect square)

Sentences are important as their variables do not need to be assigned values to examine their truth.

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg (0 \approx S(x))$ (0 is not a successor)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg (0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y ((S(x) \approx S(y)) \to (x \approx y))$ (S is one-to-one)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg(0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y((S(x) \approx S(y)) \to (x \approx y))$ (S is one-to-one)
3. $\forall x((x + 0) \approx x)$ (0 is a (right) identity for +)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg(0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y((S(x) \approx S(y)) \to (x \approx y))$ (S is one-to-one)
3. $\forall x((x + 0) \approx x)$ (0 is a (right) identity for +)
4. $\forall x \forall y((x + S(y)) \approx S(x + y))$ (+ is repeated S)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg(0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y ((S(x) \approx S(y)) \to (x \approx y))$ (S is one-to-one)
3. $\forall x ((x + 0) \approx x)$ (0 is a (right) identity for +)
4. $\forall x \forall y ((x + S(y)) \approx S(x + y))$ (+ is repeated S)
5. $\forall x ((x \cdot 0) \approx 0)$ (0 is a (right) annihilator for $\cdot$)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg (0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y ((S(x) \approx S(y)) \rightarrow (x \approx y))$ (S is one-to-one)
3. $\forall x ((x + 0) \approx x)$ (0 is a (right) identity for $+$)
4. $\forall x \forall y ((x + S(y)) \approx S(x + y))$ ($+$ is repeated S)
5. $\forall x ((x \cdot 0) \approx 0)$ (0 is a (right) annihilator for $\cdot$)
6. $\forall x \forall y ((x \cdot S(y)) \approx ((x \cdot y) + x))$ ($\cdot$ is repeated $+$)

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg(0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y((S(x) \approx S(y)) \rightarrow (x \approx y))$ (S is one-to-one)
3. $\forall x((x + 0) \approx x)$ (0 is a (right) identity for $+$)
4. $\forall x \forall y((x + S(y)) \approx S(x + y))$ ($+$ is repeated S)
5. $\forall x((x \cdot 0) \approx 0)$ (0 is a (right) annihilator for $\cdot$)
6. $\forall x \forall y((x \cdot S(y)) \approx ((x \cdot y) + x))$ ($\cdot$ is repeated $+$)
7. $((\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(S(x)))) \rightarrow \forall x \varphi(x))$ (induction for $\varphi(x)$) for each wff $\varphi(x)$.

Peano Arithmetic

As an illustrative and important example: **Peano Arithmetic** (PA) is the set of sentences

1. $\forall x \neg(0 \approx S(x))$ (0 is not a successor)
2. $\forall x \forall y((S(x) \approx S(y)) \to (x \approx y))$ (S is one-to-one)
3. $\forall x((x + 0) \approx x)$ (0 is a (right) identity for $+$)
4. $\forall x \forall y((x + S(y)) \approx S(x + y))$ ($+$ is repeated S)
5. $\forall x((x \cdot 0) \approx 0)$ (0 is a (right) annihilator for $\cdot$)
6. $\forall x \forall y((x \cdot S(y)) \approx ((x \cdot y) + x))$ ($\cdot$ is repeated $+$)
7. $((\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(S(x)))) \to \forall x \varphi(x))$ (induction for $\varphi(x)$) for each wff $\varphi(x)$.

This will be our theory of arithmetic.

Provability

Let Φ be a set of sentences (*hypotheses*) and φ a sentence.

Definition

A Φ-**proof**, or a **proof with hypotheses from** Φ, is a finite sequence $\psi_1, \ldots, \psi_n$ of wffs such that for each i, either

1. ψ_i is a hypothesis,
2. ψ_i is an *axiom of logic*, or
3. there is $j, k < i$ such that $\psi_k = (\psi_j \rightarrow \psi_i)$ (i.e. an *application of Modus Ponens* to previous steps).

Provability

Let Φ be a set of sentences (*hypotheses*) and φ a sentence.

Definition

A **Φ-proof**, or a **proof with hypotheses from Φ**, is a finite sequence $\psi_1, \ldots, \psi_n$ of wffs such that for each i, either

1. ψ_i is a hypothesis,
2. ψ_i is an *axiom of logic*, or
3. there is $j, k < i$ such that $\psi_k = (\psi_j \to \psi_i)$ (i.e. an *application of Modus Ponens* to previous steps).

Definition

φ is **provable from Φ** or Φ **proves** φ

$$\Phi \vdash \varphi$$

if there exists a Φ-proof $\psi_1, \ldots, \psi_n$ such that $\varphi = \psi_n$.

'True' Sentences of Arithmetic and Implications to PA

Definition

A sentence is **true** if it is true when interpreted in the *standard model*, i.e. when

1. quantifiers range over $\mathbb{N}$ and
2. $0, S, +, \cdot$ are the usual zero, successor, addition, and multiplication operations on $\mathbb{N}$.

'True' Sentences of Arithmetic and Implications to PA

Definition

A sentence is **true** if it is true when interpreted in the *standard model*, i.e. when

1. quantifiers range over $\mathbb{N}$ and
2. $0, S, +, \cdot$ are the usual zero, successor, addition, and multiplication operations on $\mathbb{N}$.

Proposition

If PA $\vdash \varphi$, *then* φ *is true.*

'True' Sentences of Arithmetic and Implications to PA

Definition

A sentence is **true** if it is true when interpreted in the *standard model*, i.e. when

1. quantifiers range over $\mathbb{N}$ and
2. $0, S, +, \cdot$ are the usual zero, successor, addition, and multiplication operations on $\mathbb{N}$.

Proposition

If $\mathrm{PA} \vdash \varphi$, *then* φ *is true.*

Gödel's First Incompleteness Theorem states the converse does not hold.

‘True’ Sentences of Arithmetic and Implications to PA

Definition

A sentence is **true** if it is true when interpreted in the *standard model*, i.e. when

1. quantifiers range over $\mathbb{N}$ and
2. $0, S, +, \cdot$ are the usual zero, successor, addition, and multiplication operations on $\mathbb{N}$.

Proposition

If PA $\vdash \varphi$*, then* φ *is true.*

Gödel’s First Incompleteness Theorem states the converse does not hold.

$\implies$ PA is not strong enough to capture everything about $\mathbb{N}$.

1 Meaning of Gödel's First Incompleteness Theorem

2 Proving Gödel's First Incompleteness Theorem

3 Generalizing Gödel's First Incompleteness Theorem

Arithmetization of Logic - Gödel numbers of Terms

To give a proof, we internalize our logic into arithmetic.

Arithmetization of Logic - Gödel numbers of Terms

To give a proof, we internalize our logic into arithmetic.
We recursively assign to each term t a natural number $\#(t)$ as follows:

$$\#(x_i) := 2^0 \cdot 3^0 \cdot 5^i$$

$$\#(0) := 2^0 \cdot 3^1 \qquad \#(S(t_1)) := 2^0 \cdot 3^2 \cdot 5^{\#(t_1)}$$

$$\#((t_1 + t_2)) := 2^0 \cdot 3^3 \cdot 5^{\#(t_1)} 7^{\#(t_2)} \qquad \#((t_1 \cdot t_2)) := 2^0 \cdot 3^4 \cdot 5^{\#(t_1)} 7^{\#(t_2)}$$

Arithmetization of Logic - Gödel numbers of Terms

To give a proof, we internalize our logic into arithmetic.
We recursively assign to each term t a natural number $\#(t)$ as follows:

$$\#(x_i) := 2^0 \cdot 3^0 \cdot 5^i$$

$$\#(0) := 2^0 \cdot 3^1 \qquad \#(S(t_1)) := 2^0 \cdot 3^2 \cdot 5^{\#(t_1)}$$

$$\#((t_1 + t_2)) := 2^0 \cdot 3^3 \cdot 5^{\#(t_1)} 7^{\#(t_2)} \qquad \#((t_1 \cdot t_2)) := 2^0 \cdot 3^4 \cdot 5^{\#(t_1)} 7^{\#(t_2)}$$

Definition

For a term t, $\#(t)$ is the **Gödel number** of t.

Arithmetization of Logic - Gödel numbers of Terms

To give a proof, we internalize our logic into arithmetic.
We recursively assign to each term t a natural number $\#(t)$ as follows:

$$\#(x_i) := 2^0 \cdot 3^0 \cdot 5^i$$

$$\#(0) := 2^0 \cdot 3^1 \qquad \#(S(t_1)) := 2^0 \cdot 3^2 \cdot 5^{\#(t_1)}$$

$$\#((t_1 + t_2)) := 2^0 \cdot 3^3 \cdot 5^{\#(t_1)} 7^{\#(t_2)} \qquad \#((t_1 \cdot t_2)) := 2^0 \cdot 3^4 \cdot 5^{\#(t_1)} 7^{\#(t_2)}$$

Definition

For a term t, $\#(t)$ is the **Gödel number** of t.

E.g., the Gödel number of the term

$$(x_1 \cdot (x_2 + x_3))$$

is

$$2^0 \cdot 3^4 \cdot 5^{2^0 \cdot 3^0 \cdot 5^1} \cdot 7^{2^0 \cdot 3^3 \cdot 5^{2^0 \cdot 3^0 \cdot 5^2} \cdot 7^{2^0 \cdot 3^0 \cdot 5^3}} = 3^4 \cdot 5^5 \cdot 7^{3^3 \cdot 5^{25} \cdot 7^{125}}$$

Arithmetization of Logic - Gödel numbers of Wffs

Likewise, to each wff φ:

$$\begin{aligned}
\#((t_1 \approx t_2)) &:= 2^1 \cdot 3^0 \cdot 5^{\#(t_1)} \cdot 7^{\#(t_2)} \\
\#((\psi_1 \wedge \psi_2)) &:= 2^1 \cdot 3^1 \cdot 5^{\#(\psi_1)} \cdot 7^{\#(\psi_2)} \\
&\vdots \\
\#(\neg\psi) &:= 2^1 \cdot 3^5 \cdot 5^{\#(\psi)} \\
\#(\forall x_i \psi) &:= 2^1 \cdot 3^6 \cdot 5^i \cdot 7^{\#(\psi)} \\
\#(\exists x_i \psi) &:= 2^1 \cdot 3^7 \cdot 5^i \cdot 7^{\#(\psi)} \\
\#(\bot) &:= 2^1 \cdot 3^8 \\
\#(\top) &:= 2^1 \cdot 3^9
\end{aligned}$$

Arithmetization of Logic - Gödel numbers of Wffs

Likewise, to each wff φ:

$$
\begin{aligned}
\#((t_1 \approx t_2)) &:= 2^1 \cdot 3^0 \cdot 5^{\#(t_1)} \cdot 7^{\#(t_2)} \\
\#((\psi_1 \wedge \psi_2)) &:= 2^1 \cdot 3^1 \cdot 5^{\#(\psi_1)} \cdot 7^{\#(\psi_2)} \\
&\vdots \\
\#(\neg\psi) &:= 2^1 \cdot 3^5 \cdot 5^{\#(\psi)} \\
\#(\forall x_i \psi) &:= 2^1 \cdot 3^6 \cdot 5^i \cdot 7^{\#(\psi)} \\
\#(\exists x_i \psi) &:= 2^1 \cdot 3^7 \cdot 5^i \cdot 7^{\#(\psi)} \\
\#(\bot) &:= 2^1 \cdot 3^8 \\
\#(\top) &:= 2^1 \cdot 3^9
\end{aligned}
$$

Definition

For a wff φ, $\#(\varphi)$ is the **Gödel number of** φ.

Arithmetization of Logic - Gödel numbers of proofs

Given a finite sequence

$$\psi_1, \dots, \psi_n$$

of wffs, we define

$$\#(\psi_1, \dots, \psi_n) := p_0^2 \cdot p_1^{\#(\psi_1)} \cdot p_2^{\#(\psi_2)} \cdots p_n^{\#(\psi_n)}$$

where

$$2 = p_0 < 3 = p_1 < 5 = p_2 < \cdots < p_n$$

are the first $n + 1$ prime numbers.

Arithmetization of Logic - Gödel numbers of proofs

Given a finite sequence

$$\psi_1, \ldots, \psi_n$$

of wffs, we define

$$\#(\psi_1, \ldots, \psi_n) := p_0^2 \cdot p_1^{\#(\psi_1)} \cdot p_2^{\#(\psi_2)} \cdots p_n^{\#(\psi_n)}$$

where

$$2 = p_0 < 3 = p_1 < 5 = p_2 < \cdots < p_n$$

are the first $n+1$ prime numbers.

Definition

Given a finite seuqence $\psi_1, \ldots, \psi_n$, $\#(\psi_1, \ldots, \psi_n)$ is the **Gödel number** of $\psi_1, \ldots, \psi_n$.

Arithmetization of Logic - Gödel numbers of proofs

Given a finite sequence

$$\psi_1, \ldots, \psi_n$$

of wffs, we define

$$\#(\psi_1, \ldots, \psi_n) := p_0^2 \cdot p_1^{\#(\psi_1)} \cdot p_2^{\#(\psi_2)} \cdots p_n^{\#(\psi_n)}$$

where

$$2 = p_0 < 3 = p_1 < 5 = p_2 < \cdots < p_n$$

are the first $n + 1$ prime numbers.

Definition

Given a finite seuqence $\psi_1, \ldots, \psi_n$, $\#(\psi_1, \ldots, \psi_n)$ is the **Gödel number** of $\psi_1, \ldots, \psi_n$.

In particular, we may Gödel number proofs.

Provability Formula

There is a wff

$$\mathrm{Proof}(x, y) \equiv \left(\begin{array}{l} x \text{ is (the Gödel number of) a proof of} \\ \text{the sentence } \varphi \text{ with } \#(\varphi) = y \end{array} \right)$$

Provability Formula

There is a wff

$$\mathrm{Proof}(x, y) \equiv \left(\begin{array}{l} x \text{ is (the Gödel number of) a proof of} \\ \text{the sentence } \varphi \text{ with } \#(\varphi) = y \end{array} \right)$$

allowing us to define

$$\begin{aligned} \mathrm{Provable}(y) &:= \exists x \mathrm{Proof}(x, y) \\ &\equiv (\text{there exists a proof of } \varphi, \text{ where } \#(\varphi) = y) \\ &\equiv (\varphi \text{ is provable from PA, where } \#(\varphi) = y) \end{aligned}$$

Diagonalization

Theorem (Diagonal Lemma)

Suppose T proves a sufficient fragment of PA . *If $\varphi(x)$ is a wff in the language of arithmetic, then there exists a sentence ψ such that*

$$T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$$

(The Diagonal Lemma is *constructive*: it actually gives a method for constructing ψ.)

Diagonalization

Theorem (Diagonal Lemma)

Suppose T proves a sufficient fragment of PA *. If $\varphi(x)$ is a wff in the language of arithmetic, then there exists a sentence ψ such that*

$$T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$$

(The Diagonal Lemma is *constructive*: it actually gives a method for constructing ψ.)

Definition

G, the **Gödel sentence**, is the sentence satisfying

$$\mathrm{PA} \vdash (G \leftrightarrow \neg \mathrm{Provable}(\#(G)))$$

Diagonalization

Theorem (Diagonal Lemma)

Suppose T proves a sufficient fragment of PA . If $\varphi(x)$ is a wff in the language of arithmetic, then there exists a sentence ψ such that

$$T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$$

(The Diagonal Lemma is *constructive*: it actually gives a method for constructing ψ.)

Definition

G, the **Gödel sentence**, is the sentence satisfying

$$\text{PA} \vdash (G \leftrightarrow \neg\text{Provable}(\#(G)))$$

In other words,

$$G \equiv (\text{I am not provable})$$

Truth and Unprovability of the Gödel Sentence

By definition, PA $\vdash (G \leftrightarrow \neg\mathrm{Provable}(\#(G)))$.

Unprovable: Suppose G were provable, with $\psi_1, \ldots, \psi_n$ a proof. Then $\#(\psi_1, \ldots, \psi_n)$ realizes that $\mathrm{Provable}(\#(G))$ is true and hence provable. But then $\neg G$ is provable, a contradiction.

Truth and Unprovability of the Gödel Sentence

By definition, PA $\vdash (G \leftrightarrow \neg\mathrm{Provable}(\#(G)))$.

Unprovable: Suppose G were provable, with $\psi_1, \ldots, \psi_n$ a proof. Then $\#(\psi_1, \ldots, \psi_n)$ realizes that $\mathrm{Provable}(\#(G))$ is true and hence provable. But then $\neg G$ is provable, a contradiction.

True: If G was not true when interpreted in $\mathbb{N}$, then it cannot be provable, so

$$\neg\mathrm{Prove}(n, \#(G)) \quad \text{for each } n \in \mathbb{N}.$$

is true for each n. Thus,

$$\forall n \neg\mathrm{Proof}(n, \#(G))$$

is true, which is equivalent to G. Contradiction.

Looking Forward

If G is true but unprovable, why not accept

$$T := \mathrm{PA} \cup \{G\}$$

as our new theory of arithmetic?

Looking Forward

If G is true but unprovable, why not accept

$$T := \mathrm{PA} \cup \{G\}$$

as our new theory of arithmetic?
Our method of proving Gödel's First Incompleteness Theorem would work to show that there is a new sentence G' which is true but unprovable from $T := \mathrm{PA} \cup \{G\}$.

Looking Forward

If G is true but unprovable, why not accept

$$T := \mathrm{PA} \cup \{G\}$$

as our new theory of arithmetic?

Our method of proving Gödel's First Incompleteness Theorem would work to show that there is a new sentence G' which is true but unprovable from $T := \mathrm{PA} \cup \{G\}$.

Likewise, there is a sentence G'' which is true but unprovable from $T := \mathrm{PA} \cup \{G, G'\}$, and so on.

Looking Forward

If G is true but unprovable, why not accept

$$T := \mathrm{PA} \cup \{G\}$$

as our new theory of arithmetic?

Our method of proving Gödel's First Incompleteness Theorem would work to show that there is a new sentence G' which is true but unprovable from $T := \mathrm{PA} \cup \{G\}$.

Likewise, there is a sentence G'' which is true but unprovable from $T := \mathrm{PA} \cup \{G, G'\}$, and so on.

Even if we considered $T := \mathrm{PA} \cup \{G, G', G'', \ldots\}$, there would exist a $\tilde{G}$ which is true but unprovable from T.

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

What were the essential ingredients in our proof in the case of PA?

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

What were the essential ingredients in our proof in the case of PA?

1. Existence of $\mathrm{Provable}(x)$ (to encode provability in arithmetic)

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

What were the essential ingredients in our proof in the case of PA?

1. Existence of $\mathrm{Provable}(x)$ (to encode provability in arithmetic)
2. Diagonal Lemma (to build the Gödel sentence G)

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

What were the essential ingredients in our proof in the case of PA?

1. Existence of $\mathrm{Provable}(x)$ (to encode provability in arithmetic)
2. Diagonal Lemma (to build the Gödel sentence G)
3. PA does not prove a contradiction (to show G is not provable)

Analysis of Our Proof Sketch

Want to repeat the proof of Gödel's First Incompleteness Theorem in more general theories of arithmetic. Let T be a theory (set of sentences) of arithmetic.

What were the essential ingredients in our proof in the case of PA?

1. Existence of $\mathrm{Provable}(x)$ (to encode provability in arithmetic)
2. Diagonal Lemma (to build the Gödel sentence G)
3. PA does not prove a contradiction (to show G is not provable)
4. If PA does not simultaneously prove $\neg\varphi(n)$ for every $n \in \mathbb{N}$ and $\exists x \varphi(x)$ (to show $\neg G$ is not provable and G is true)

Primitive-Recursiveness

Definition

A function $f : \mathbb{N}^k \to \mathbb{N}$ is **primitive-recursive** if it is computable by an algorithm which does not use any unbounded searches (no *while* loops).
A subset $R \subseteq \mathbb{N}^k$ is **primitive-recursive** if the function

$$\chi_R(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } (x_1, \ldots, x_k) \in R \\ 0 & \text{otherwise} \end{cases}$$

is primitive-recursive.

Primitive-Recursiveness

Definition

A function $f : \mathbb{N}^k \to \mathbb{N}$ is **primitive-recursive** if it is computable by an algorithm which does not use any unbounded searches (no *while* loops).
A subset $R \subseteq \mathbb{N}^k$ is **primitive-recursive** if the function
$\chi_R(x_1, \ldots, x_k) = \begin{cases} 1 & \text{if } (x_1, \ldots, x_k) \in R \\ 0 & \text{otherwise} \end{cases}$ is primitive-recursive.

Proposition

If $R \subseteq \mathbb{N}^k$ is primitive-recursive, there exists a wff $\varphi(x_1, \ldots, x_k)$ such that for every $(n_1, \ldots, n_k)$

$$\mathrm{PA} \vdash \varphi(n_1, \ldots, n_k) \quad \textit{if and only if} \quad (n_1, \ldots, n_k) \in R$$

Effectively-Generated Theories

The definability of provability from PA depended on the fact that, given a natural number, we can determine whether it is a valid proof in PA .

Effectively-Generated Theories

The definability of provability from PA depended on the fact that, given a natural number, we can determine whether it is a valid proof in PA .

Definition

A theory of arithmetic T is **effectively-generated** if there is a primitive-recursive function which enumerates the (Gödel numbers of) elements of T. (Equivalently, can drop 'primitive'.)

Effectively-Generated Theories

The definability of provability from PA depended on the fact that, given a natural number, we can determine whether it is a valid proof in PA .

Definition

A theory of arithmetic T is **effectively-generated** if there is a primitive-recursive function which enumerates the (Gödel numbers of) elements of T. (Equivalently, can drop 'primitive'.)

PA is effectively-generated (our listing of its elements outlines a procedure for enumerating those elements), as are $\mathrm{PA} \cup \{G\}$, $\mathrm{PA} \cup \{G, G'\}$, etc.

Modifying our Proof

To support being only effectively-generated, we must modify our encoding of proofs:

Modifying our Proof

To support being only effectively-generated, we must modify our encoding of proofs:

Definition

Say that n is (the Gödel number of) a *T-proof* if

$$n = p_0^2 \cdot p_1^{i_1} \cdots p_m^{i_m}$$

where

1. each i_k is either the Gödel number of a formula or equal is $i_k = 2^3 \cdot 5^{j_k}$ and
2. $\psi_1, \ldots, \psi_m$ is a proof with hypotheses from T, where either $\#(\psi_k) = i_k$ or ψ_k is the j_k-th element of T.

Modifying our Proof

To support being only effectively-generated, we must modify our encoding of proofs:

Definition

Say that n is (the Gödel number of) a T*-proof* if

$$n = p_0^2 \cdot p_1^{i_1} \cdots p_m^{i_m}$$

where

1. each i_k is either the Gödel number of a formula or equal is $i_k = 2^3 \cdot 5^{j_k}$ and
2. $\psi_1, \ldots, \psi_m$ is a proof with hypotheses from T, where either $\#(\psi_k) = i_k$ or ψ_k is the j_k-th element of T.

With this modification and T being effectively-generated, the relation

$$\mathrm{Provable}_T(x) := (\text{there is a proof of } x \text{ from } T)$$

is primitive-recursive.

How much arithmetic do we need?

To construct G, it remains to show that $\mathrm{Provable}_T(x)$ can be defined by a wff.

To both finish the proof of existence of the wff $\mathrm{Provable}(x)$ as well as prove the Diagonal Lemma, we need our theory T to contain enough arithmetical truths to show that every primitive-recursive predicate is definable.

One such benchmark for this is Q, *Robinsin Arithmetic*, which drops the induction axioms from PA.

Proposition

If $R \subseteq \mathbb{N}^k$ is primitive-recursive, there exists a wff $\varphi(x_1, \ldots, x_k)$ such that for every $(n_1, \ldots, n_k)$

$$\mathrm{Q} \vdash \varphi(n_1, \ldots, n_k) \quad \textit{if and only if} \quad (n_1, \ldots, n_k) \in R$$

ω-Consistency

To conclude that neither G nor $\neg G$ are provable, we need some *consistency* properties for T.

ω-Consistency

To conclude that neither G nor $\neg G$ are provable, we need some *consistency* properties for T.

- To show G was unprovable we only needed that PA did not prove a contradiction.

ω-Consistency

To conclude that neither G nor $\neg G$ are provable, we need some *consistency* properties for T.

- To show G was unprovable we only needed that PA did not prove a contradiction.
- To show that G was true (which showed $\neg G$ was not provable) we needed something stronger: that PA could not simultaneously prove $\neg\varphi(n)$ for each $n \in \mathbb{N}$ and also prove $\exists x \varphi(x)$.

Definition

A set of sentences T is **consistent** if it does not prove a contradiction.
A set of sentences T is ω-**consistent** if it does not simultaneously prove $\neg\varphi(n)$ for each $n \in \mathbb{N}$ and $\exists x \varphi(x)$.

(Note that ω-consistency implies consistency.)

Gödel's First Incompleteness Theorem, revisited

Theorem (Gödel's First Incompleteness Theorem)

If T is an effectively-generated, ω-consistent theory in the language of arithmetic and $\mathsf{Q} \subseteq T$, then there exists a sentence G such that neither G nor $\neg G$ are provable from T, i.e. T is **incomplete**.

Rosser's Trick

We can weaken the hypothesis of ω-consistency to that of consistency by replacing $\mathrm{Provable}(x)$ with a different wff.

Rosser's Trick

We can weaken the hypothesis of ω-consistency to that of consistency by replacing $\mathrm{Provable}(x)$ with a different wff.
Let $\mathrm{neg}(y)$ be the Gödel number of $\neg\varphi$, where $\#(\varphi) = y$.
Define

$$\mathrm{Proof}_T^R(x,y) := \mathrm{Proof}_T(x,y) \wedge \neg\exists z(z \leq x \wedge \mathrm{Proof}_T(z, \mathrm{neg}(y)))$$

$$\equiv \left(\begin{array}{l} x \text{ encodes a proof of } \varphi \text{ and} \\ \text{there is no shorter proof of } \neg\varphi, \\ \text{where } \#(\varphi) = y \end{array} \right)$$

and

$$\mathrm{Provable}_T^R(y) := \exists x \mathrm{Proof}_T^R(x,y)$$

Rosser's Trick, Continued

Using the Diagonal Lemma with $\neg\mathrm{Provable}_T^R(x)$ yields:

Definition

ρ, the **Gödel-Rosser Sentence** for T, is the sentence satisfying

$$T \vdash (\rho \leftrightarrow \neg\mathrm{Provable}_T^R(\#(\rho)))$$

Rosser's Trick, Continued

Using the Diagonal Lemma with $\neg\text{Provable}_T^R(x)$ yields:

Definition

ρ, the **Gödel-Rosser Sentence** for T, is the sentence satisfying

$$T \vdash (\rho \leftrightarrow \neg\text{Provable}_T^R(\#(\rho)))$$

T, only under the hypothesis of consistency and the other hypotheses of Gödel's First Incompleteness Theorem, neither proves ρ nor $\neg\rho$.

Rosser's Trick, Continued

Using the Diagonal Lemma with $\neg \mathrm{Provable}_T^R(x)$ yields:

Definition

ρ, the **Gödel-Rosser Sentence** for T, is the sentence satisfying

$$T \vdash (\rho \leftrightarrow \neg \mathrm{Provable}_T^R(\#(\rho)))$$

T, only under the hypothesis of consistency and the other hypotheses of Gödel's First Incompleteness Theorem, neither proves ρ nor $\neg\rho$.

Theorem (Gödel-Rosser Incompleteness Theorem)

If T is an effectively-generated, consistent theory in the language of arithmetic and $\mathrm{Q} \subseteq T$, then T is incomplete.

Thank you!

Questions?

Proof of Diagonal Lemma

Consider the primitive recursive function

$$\mathrm{diag}(n) := \begin{cases} \#(\chi(n)) & \text{if } \#(\chi) = n \\ 0 & \text{otherwise} \end{cases}$$

Since $T \supseteq \mathrm{Q}$, there exists a wff $\mathrm{Diag}(x, y)$ such that

$$\mathrm{diag}(x) = y \quad \text{if and only if} \quad T \vdash \mathrm{Diag}(x, y)$$

Let

$$\chi(x) := \exists y(\mathrm{Diag}(x, y) \wedge \varphi(y))$$

and

$$\psi := \chi(\#(\chi))$$

Claim: $T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$.

Proof of Diagonal Lemma

Claim: $T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$.

$\rightarrow$: By definition,

$$T \cup \{\psi\} \vdash \chi(\#(\chi)) \quad [\equiv \exists y(\mathrm{Diag}(\#(\chi), y) \wedge \varphi(y))]$$

But $y = \#(\psi)$ is the only number such that $T \vdash \mathrm{Diag}(\#(\chi), y)$, so

$$T \cup \{\psi\} \vdash \varphi(\#(\psi))$$

The Deduction Theorem then proves $T \vdash (\psi \rightarrow \varphi(\#(\psi)))$.

Proof of Diagonal Lemma

Claim: $T \vdash (\psi \leftrightarrow \varphi(\#(\psi)))$.

$\leftarrow$: $T \vdash \mathrm{Diag}(\#(\chi), \#(\psi))$, so

$$T \cup \{\varphi(\#(\psi))\} \vdash (\mathrm{Diag}(\#(\chi), \#(\psi)) \wedge \varphi(\#(\psi)))$$

and hence

$$T \cup \{\varphi(\#(\psi))\} \vdash \exists y(\mathrm{Diag}(\#(\chi), y) \wedge \varphi(y)) \quad [\equiv \psi]$$

The Deduction Theorem then proves $T \vdash (\varphi(\#(\psi)) \rightarrow \psi)$.

Gödel-Rosser Sentence – More Details

That T does not prove ρ is analogous to our original proof.
Now suppose for the sake of a contradiction that T proves $\neg\rho$. Let

$$e := \text{a natural number encoding a proof of } \neg\rho \text{ from } T$$

Since T is consistent, there is no code for a proof of ρ in T, so

$$\mathrm{Proof}_T^R(e, \mathrm{neg}(\#(\rho)))$$

Then T proves (since $T \supseteq \mathrm{Q}$)

$$\forall x(e \leq x \to \exists z \leq x \mathrm{Proof}_T(z, \neg(\#(\rho))))$$

and (using consistency)

$$\neg\exists x < e \mathrm{Proof}_T(x, \#(\rho))$$

Then T proves

$$\forall x(\mathrm{Proof}_T(x, \#(\rho)) \to e \leq x)$$

so that T proves

$$\forall x(\mathrm{Proof}_T(x, \#(\rho)) \to \exists z \leq x \mathrm{Proof}_T(z, \neg(\#(\rho))))$$