

Lucas Waclawczyk

Decidability of Logical Theories

Proseminar Theoretical Computer Science // Dresden, June 20, 2020

Can we mechanize mathematics?

Content

First-Order Logic

$\text{Th}(\mathbb{N}, +)$ – A Decidable Theory

$\text{Th}(\mathbb{N}, +, \times)$ – An Undecidable Theory

Gödel's Incompleteness Theorem

Conclusion

References

First-Order Logic

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg (R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall , \exists

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots
- Boolean operators: \wedge, \vee, \neg

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots
- Boolean operators: \wedge, \vee, \neg
- Relation symbols: R_1, R_2, \dots

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots
- Boolean operators: \wedge, \vee, \neg
- Relation symbols: R_1, R_2, \dots
- Special characters: $[,], (,)$

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots
- Boolean operators: \wedge, \vee, \neg
- Relation symbols: R_1, R_2, \dots
- Special characters: $[,], (,)$

Simplified here:

- mostly "sentences"
(no free variables)

Syntax

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Formulas consist of:

- Quantifiers: \forall, \exists
- Variables: p, q, x, y, \dots
- Boolean operators: \wedge, \vee, \neg
- Relation symbols: R_1, R_2, \dots
- Special characters: $[,], (,)$

Simplified here:

- mostly "sentences"
(no free variables)
- only prenex normal form
(all quantifiers on the left)

Semantics

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

Semantics

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

$\stackrel{!}{=} \text{"There are infinitely many prime numbers."}$

Semantics

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$

-
- *universe \mathcal{U}*
 - possible values for variables
 - here \mathbb{N}

Semantics

$$\forall q \exists p \forall x \forall y \left[R_1(p, q) \wedge \left(\neg(R_1(x, 1) \wedge R_1(y, 1)) \vee R_2(x, y, p) \right) \right]$$
$$= \forall q \exists p \forall x, y [p > q \wedge (x, y > 1 \rightarrow xy \neq p)]$$

-
- *universe \mathcal{U}*
 - possible values for variables
 - here \mathbb{N}
 - *model \mathcal{M}*
 - universe + assignment of relations
 - here $(\mathbb{N}, >_2, (\times \neq)_3)$

Semantics

$$\forall q \exists p \forall x, y [p > q \wedge (x, y > 1 \rightarrow xy \neq p)]$$

= "There are infinitely many prime numbers."

- *universe \mathcal{U}*
 - possible values for variables
 - here \mathbb{N}
- *model \mathcal{M}*
 - universe + assignment of relations
 - here $(\mathbb{N}, >_2, (\times \neq)_3)$

Semantics

$$\forall q \exists p \forall x, y [p > q \wedge (x, y > 1 \rightarrow xy \neq p)]$$

$$\text{vs. } [(x + x = y) \vee (x \geq y)]$$

-
- *universe* \mathcal{U}
 - possible values for variables
 - here \mathbb{N}
 - *model* \mathcal{M}
 - universe + assignment of relations
 - here $(\mathbb{N}, >_2, (\times \neq)_3)$
 - *language* $L(\mathcal{M})$
 - sentences that make sense in \mathcal{M}
 - here $L(\mathbb{N}, >_2, (\times \neq)_3)$

Semantics

$$\forall q \exists p \forall x, y [p > q \wedge (x, y > 1 \rightarrow xy \neq p)]$$

$$\text{vs. } \forall y \exists x [\neg (xx \neq y)]$$

-
- *universe* \mathcal{U}
 - possible values for variables
 - here \mathbb{N}
 - *model* \mathcal{M}
 - universe + assignment of relations
 - here $(\mathbb{N}, >_2, (\times \neq)_3)$
 - *language* $L(\mathcal{M})$
 - sentences that make sense in \mathcal{M}
 - here $L(\mathbb{N}, >_2, (\times \neq)_3)$
 - *theory* $\text{Th}(\mathcal{M})$
 - true sentences formed with \mathcal{M}
 - here $\text{Th}(\mathbb{N}, >_2, (\times \neq)_3)$

What Is Decidability?

- here: for logic (there is a more general definition)
- let \mathcal{M} be a model, $\varphi \in L(\mathcal{M})$

$\text{Th}(\mathcal{M})$ decidable $:=$

there is an algorithm that decides whether φ is true in \mathcal{M}

$\text{Th}(\mathbb{N}, +)$ – A Decidable Theory

Theorem 1

Theorem

$Th(\mathbb{N}, +)$ is decidable.

Theorem 1

Theorem

$Th(\mathbb{N}, +)$ is decidable.

i.e., there is an algorithm that can decide,
whether a sentence $\varphi \in L(\mathbb{N}, +)$ is true or false.

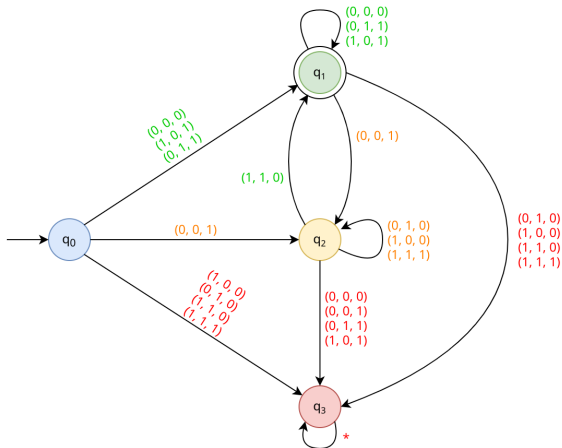
Proof of Theorem 1

Idea: Construct an automaton that accepts an (almost) empty input iff the given sentence is true.

Let $i \in \mathbb{N} \setminus \{0\}$ and define $\Sigma_i := \{0, 1\}^i$ and $\Sigma_0 := \{()\}$.
An example for a word in Σ_3 :

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \sim \begin{pmatrix} 3 \\ 9 \\ 12 \end{pmatrix}$$

Review: Automata



Example automaton that accepts $+_3$ encoded in Σ_3

Proof of Theorem 1

Now, let

- $k \in \{0, \dots, l\}$
- $\varphi = Q_1 x_1 \dots Q_l x_l [\psi(x_1, \dots, x_l)] \in L(\mathbb{N}, +)$ where $Q_1, \dots, Q_l \in \{\forall, \exists\}$

Proof of Theorem 1

Now, let

- $k \in \{0, \dots, l\}$
- $\varphi = Q_1 x_1 \dots Q_l x_l [\psi(x_1, \dots, x_l)] \in L(\mathbb{N}, +)$ where $Q_1, \dots, Q_l \in \{\forall, \exists\}$
- $\varphi_k := Q_{k+1} x_{k+1} \dots Q_l x_l [\psi(x_1, \dots, x_l)]$
 - $\Rightarrow \varphi_0 = \varphi$
 - $\Rightarrow \varphi_l = \psi(x_1, \dots, x_l)$

Proof of Theorem 1

Now, let

- $k \in \{0, \dots, l\}$
- $\varphi = Q_1 x_1 \dots Q_l x_l [\psi(x_1, \dots, x_l)] \in L(\mathbb{N}, +)$ where $Q_1, \dots, Q_l \in \{\forall, \exists\}$
- $\varphi_k := Q_{k+1} x_{k+1} \dots Q_l x_l [\psi(x_1, \dots, x_l)]$
 $\Rightarrow \varphi_0 = \varphi$
 $\Rightarrow \varphi_l = \psi(x_1, \dots, x_l)$
- $\varphi_k(a_1, \dots, a_k) := Q_{k+1} x_{k+1} \dots Q_l x_l [\psi(a_1, \dots, a_k, x_{k+1}, \dots, x_l)]$
where $a_1, \dots, a_k \in \mathbb{N}$

Proof of Theorem 1

Now, let

- $k \in \{0, \dots, l\}$
- $\varphi = Q_1 x_1 \dots Q_l x_l [\psi(x_1, \dots, x_l)] \in L(\mathbb{N}, +)$ where $Q_1, \dots, Q_l \in \{\forall, \exists\}$
- $\varphi_k := Q_{k+1} x_{k+1} \dots Q_l x_l [\psi(x_1, \dots, x_l)]$
 $\Rightarrow \varphi_0 = \varphi$
 $\Rightarrow \varphi_l = \psi(x_1, \dots, x_l)$
- $\varphi_k(a_1, \dots, a_k) := Q_{k+1} x_{k+1} \dots Q_l x_l [\psi(a_1, \dots, a_k, x_{k+1}, \dots, x_l)]$
where $a_1, \dots, a_k \in \mathbb{N}$

$\Rightarrow \varphi_l$ is only a Boolean expression.

Proof of Theorem 1

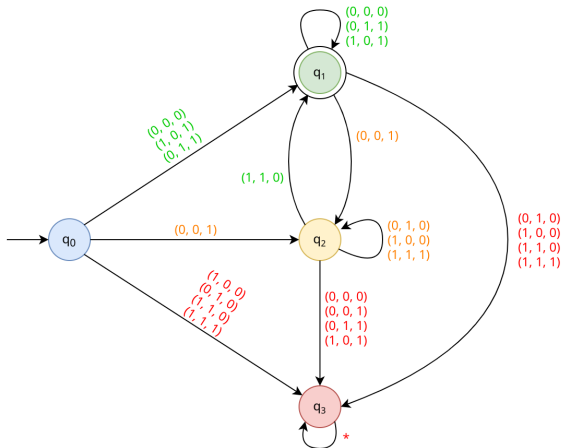
Construct an automaton A_I that behaves like $\varphi_I = \psi$, meaning A_I accepts exactly the tuples $(a_1, \dots, a_I) \in \mathbb{N}^I$ for which $\varphi_I(a_1, \dots, a_I)$ is true:

Proof of Theorem 1

Construct an automaton A_I that behaves like $\varphi_I = \psi$, meaning A_I accepts exactly the tuples $(a_1, \dots, a_I) \in \mathbb{N}^I$ for which $\varphi_I(a_1, \dots, a_I)$ is true:

- Take one addition automaton for each addition term in φ_I

Proof of Theorem 1



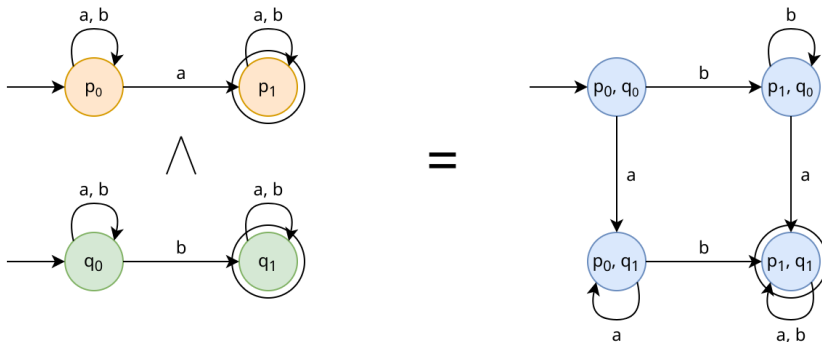
Addition automaton

Proof of Theorem 1

Construct an automaton A_I that behaves like $\varphi_I = \psi$, meaning A_I accepts exactly the tuples $(a_1, \dots, a_I) \in \mathbb{N}^I$ for which $\varphi_I(a_1, \dots, a_I)$ is true:

- Take one addition automaton for each addition term in φ_I
 - Combine them:
 - automaton product for \wedge
 - automaton union for \vee
 - automaton complement for \neg
- in a way that they behave like φ_I .

Proof of Theorem 1



Conjunction of two simple automata

Proof of Theorem 1

Construct an automaton A_I that behaves like $\varphi_I = \psi$, meaning A_I accepts exactly the tuples $(a_1, \dots, a_I) \in \mathbb{N}^I$ for which $\varphi_I(a_1, \dots, a_I)$ is true:

- Take one addition automaton for each addition term in φ_I
- Combine them:
 - automaton product for \wedge
 - automaton union for \vee
 - automaton complement for \negin a way that they behave like φ_I .

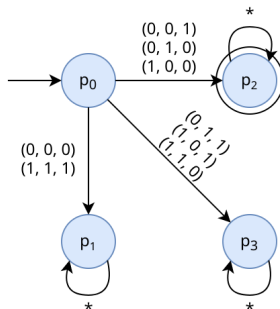
Important: There is an algorithm that constructs A_I from φ_I .

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

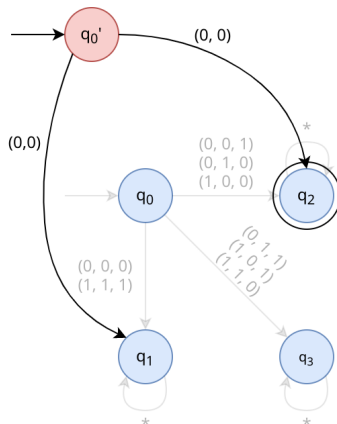
- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

Proof of Theorem 1



Example construction of non-deterministic guessing

Proof of Theorem 1



Example construction of non-deterministic guessing

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

If however $Q_I = \forall$, use complementation twice ($\forall x_I \varphi_I = \neg \exists x_I \neg \varphi_I$)

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

If however $Q_I = \forall$, use complementation twice ($\forall x_I \varphi_I = \neg \exists x_I \neg \varphi_I$)

\Rightarrow Inductively construct A_{I-2}, \dots, A_0

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

If however $Q_I = \forall$, use complementation twice ($\forall x_I \varphi_I = \neg \exists x_I \neg \varphi_I$)

\Rightarrow Inductively construct A_{I-2}, \dots, A_0

$\Rightarrow A_k$ accepts input $(a_1, \dots, a_k) \in \mathbb{N}^k \Leftrightarrow \varphi_k(a_1, \dots, a_k)$ is true

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

If however $Q_I = \forall$, use complementation twice ($\forall x_I \varphi_I = \neg \exists x_I \neg \varphi_I$)

\Rightarrow Inductively construct A_{I-2}, \dots, A_0

$\Rightarrow A_k$ accepts input $(a_1, \dots, a_k) \in \mathbb{N}^k \Leftrightarrow \varphi_k(a_1, \dots, a_k)$ is true

$\Rightarrow A_0$ accepts input $() \Leftrightarrow \varphi_0 = \varphi$ is true

Proof of Theorem 1

If $Q_I = \exists$, construct automaton A_{I-1} from A_I by

- copying states of A_I
- adding a new starting state
- making A_{I-1} guess the right a_I non-deterministically

If however $Q_I = \forall$, use complementation twice ($\forall x_I \varphi_I = \neg \exists x_I \neg \varphi_I$)

\Rightarrow Inductively construct A_{I-2}, \dots, A_0

$\Rightarrow A_k$ accepts input $(a_1, \dots, a_k) \in \mathbb{N}^k \Leftrightarrow \varphi_k(a_1, \dots, a_k)$ is true

$\Rightarrow A_0$ accepts input $() \Leftrightarrow \varphi_0 = \varphi$ is true

Let the algorithm return " $\varphi \in \text{Th}(\mathbb{N}, +)$ " $\Leftrightarrow A_0$ accepts input $()$

$\text{Th}(\mathbb{N}, +, \times)$ – An Undecidable Theory

Theorem 2

Theorem

$Th(\mathbb{N}, +, \times)$ is undecidable.

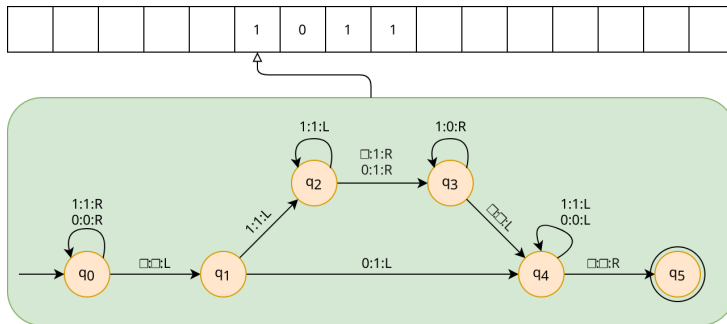
Theorem 2

Theorem

$Th(\mathbb{N}, +, \times)$ is undecidable.

i.e., there is no algorithm that can decide,
whether a sentence $\varphi \in L(\mathbb{N}, +, \times)$ is true or false.

Turing Machines



Example turing machine
adapted from: [2]

Proof Idea for Theorem 2

- **The word problem for Turing machines is undecidable.** ([5], Satz 5.6)

Proof Idea for Theorem 2

- The word problem for Turing machines is undecidable. ([5], Satz 5.6)
- There is a mapping reduction that translates
 - a Turing machine M and a string w

Proof Idea for Theorem 2

- The word problem for Turing machines is undecidable. ([5], Satz 5.6)
- There is a mapping reduction that translates
 - a Turing machine M and a string w
 - to a formula $\varphi_{M,w}(x) \in L(\mathbb{N}, +, \times)$ that contains only one free variable x , such that

Proof Idea for Theorem 2

- The word problem for Turing machines is undecidable. ([5], Satz 5.6)
- There is a mapping reduction that translates
 - a Turing machine M and a string w
 - to a formula $\varphi_{M,w}(x) \in L(\mathbb{N}, +, \times)$ that contains only one free variable x , such that
 - $\varphi_{M,w}(x)$ is true $\Leftrightarrow M$ accepts w (x = computation history suitably encoded)

Proof Idea for Theorem 2

- The word problem for Turing machines is undecidable. ([5], Satz 5.6)
- There is a mapping reduction that translates
 - a Turing machine M and a string w
 - to a formula $\varphi_{M,w}(x) \in L(\mathbb{N}, +, \times)$ that contains only one free variable x , such that
 - $\varphi_{M,w}(x)$ is true $\Leftrightarrow M$ accepts w (x = computation history suitably encoded)

Assume $\text{Th}(\mathbb{N}, +, \times)$ is decidable.

\Rightarrow The formulas $\exists x [\varphi_{M,w}(x)] \in L(\mathbb{N}, +, \times)$ are decidable.

Proof Idea for Theorem 2

- The word problem for Turing machines is undecidable. ([5], Satz 5.6)
- There is a mapping reduction that translates
 - a Turing machine M and a string w
 - to a formula $\varphi_{M,w}(x) \in L(\mathbb{N}, +, \times)$ that contains only one free variable x , such that
 - $\varphi_{M,w}(x)$ is true $\Leftrightarrow M$ accepts w (x = computation history suitably encoded)

Assume $\text{Th}(\mathbb{N}, +, \times)$ is decidable.

- \Rightarrow The formulas $\exists x [\varphi_{M,w}(x)] \in L(\mathbb{N}, +, \times)$ are decidable.
- \Rightarrow The word problem for Turing machines is decidable.

Proof Idea for Theorem 2

- **The word problem for Turing machines is undecidable.** ([5], Satz 5.6) ⚡
- There is a mapping reduction that translates
 - a Turing machine M and a string w
 - to a formula $\varphi_{M,w}(x) \in L(\mathbb{N}, +, \times)$ that contains only one free variable x , such that
 - $\varphi_{M,w}(x)$ is true $\Leftrightarrow M$ accepts w (x = computation history suitably encoded)

Assume $\text{Th}(\mathbb{N}, +, \times)$ is decidable.

- \Rightarrow The formulas $\exists x [\varphi_{M,w}(x)] \in L(\mathbb{N}, +, \times)$ are decidable.
- \Rightarrow The word problem for Turing machines is decidable.
- \Rightarrow ⚡

Thinking further...

Is there a true, unprovable sentence?

Thinking further...

Is there a true, unprovable sentence?

A proof is a series of implications and can be written as a string over some alphabet (here: $L(\mathbb{N}, +, \times)$). Assumptions:

A_1 Proofs can be checked by a machine.

A_2 Provable statements are true.

Thinking further...

Is there a true, unprovable sentence?

A proof is a series of implications and can be written as a string over some alphabet (here: $L(\mathbb{N}, +, \times)$). Assumptions:

A_1 Proofs can be checked by a machine.

A_2 Provable statements are true.

Lemmas:

1. The provable statements of $\text{Th}(\mathbb{N}, +, \times)$ are Turing recognizable.

Proof idea: Just try all possible (suitably encoded) proofs.

Thinking further...

Is there a true, unprovable sentence?

A proof is a series of implications and can be written as a string over some alphabet (here: $L(\mathbb{N}, +, \times)$). Assumptions:

A_1 Proofs can be checked by a machine.

A_2 Provable statements are true.

Lemmas:

1. The provable statements of $\text{Th}(\mathbb{N}, +, \times)$ are Turing recognizable.

Proof idea: Just try all possible (suitably encoded) proofs.

2. There is a (true) statement in $\text{Th}(\mathbb{N}, +, \times)$ that is not provable.

Proof idea: Contradiction to Theorem 2 by using 1.

Gödel's Incompleteness Theorem



A True, Unprovable Statement

Theorem

We can construct a true statement in $Th(\mathbb{N}, +, \times)$, that is not provable.

A True, Unprovable Statement

Theorem

We can construct a true statement in $Th(\mathbb{N}, +, \times)$, that is not provable.

Proof: Let M be a Turing machine that operates as follows.

- Delete the input
- Look for proof of $\varphi := \neg \exists x [\varphi_{M,0}(x)] \in Th(\mathbb{N}, +, \times)$
- Go to final state $:\Leftrightarrow$ proof for φ has been found

A True, Unprovable Statement

Theorem

We can construct a true statement in $Th(\mathbb{N}, +, \times)$, that is not provable.

Proof: Let M be a Turing machine that operates as follows.

- Delete the input
- Look for proof of $\varphi := \neg \exists x [\varphi_{M,0}(x)] \in Th(\mathbb{N}, +, \times)$
- Go to final state $:\Leftrightarrow$ proof for φ has been found

$\Rightarrow \varphi$ is the wanted statement.

Gödel's Method

- Construct the statement "This statement cannot be proved by the axioms."

Gödel's Method

- Construct the statement "This statement cannot be proved by the axioms."
- Argue against just adding this statement to the axiom.

Gödel's Method

- Construct the statement "This statement cannot be proved by the axioms."
- Argue against just adding this statement to the axiom.

\Rightarrow Incompleteness 😊

Conclusion

Conclusion

\Rightarrow There are (very simple) undecidable logical theories.

Conclusion

- \Rightarrow There are (very simple) undecidable logical theories.
- \Rightarrow No sound arithmetical system can be complete.

Conclusion

- ⇒ There are (very simple) undecidable logical theories.
- ⇒ No sound arithmetical system can be complete.
- ⇒ Mathematics cannot be mechanized.

References

References

- [1] Martin Alessandro - Own work, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=75082873>
- [2] https://www.inf-schule.de/grenzen/berechenbarkeit/turingmaschine/station_turingmaschine
- [3] Michael Sipser: Introduction to the Theory of Computation. Thomson Course Technology, 2006
- [4] <https://www.youtube.com/watch?v=04ndIDcDSGc>
- [5] Prof. Dr. Franz Baader: Skript Theoretische Informatik und Logik (Sommersemester 2020), TU Dresden