Epistemic Arithmetic

Eric Pacuit

University of Maryland
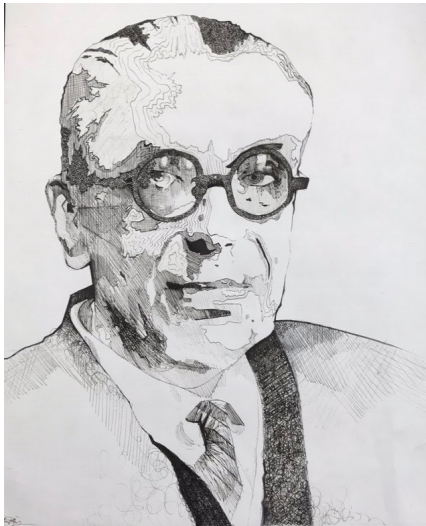
Lecture 1, ESSLLI 2025

July 28, 2025
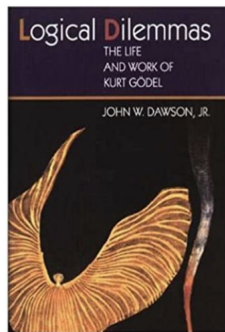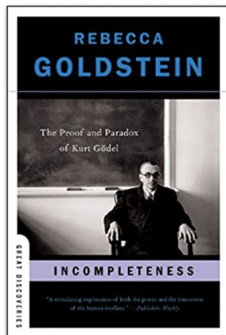
Plan

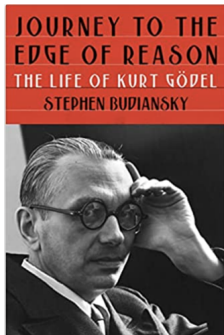- Introduction: Smullyan's Machine
- Background
  - Formal Arithmetic
  - Gödel's Incompleteness Theorems
  - Names and Gödel numbering
  - Fixed Point Theorem
- Provability predicate and Löb's Theorem
- Provability logic
- Predicate approach to modality
- A Primer on Epistemic and Doxastic Logic
- Anti-Expert Paradoxes
- The Knower Paradox and variants
- Epistemic Arithmetic
- Gödel's Disjunction

Introduction

Kurt Gödel (1906 - 1978)

plato.stanford.edu/entries/goedel/

JOURNEY TO THE
EDGE OF REASON
THE LIFE OF KURT GÖDEL
STEPHEN BUDIANSKY

REBECCA
GOLDSTEIN

The Proof and Paradox
of Kurt Gödel

INCOMPLETENESS

"A stimulating exploration of both the genius and the limitations
of our human intellect." —Publishers Weekly

Logical Dilemmas
THE LIFE
AND WORK OF
KURT GÖDEL

JOHN W. DAWSON, JR.

1929   Completeness of First-Order Logic

1931   First and Second Incompleteness Theorems

1933   Translation of classical logic in intuitionistic logic

1936   Speed-up Theorems

1938   Consistency of the Continuum Hypothesis

1949   Work on General Relativity

1958   The "Dialectica interpretation"

1929   Completeness of First-Order Logic

1931   First and Second Incompleteness Theorems

1933   Translation of classical logic in intuitionistic logic

1936   Speed-up Theorems

1938   Consistency of the Continuum Hypothesis

1949   Work on General Relativity

1958   The "Dialectica interpretation"

Kurt Gödel
ON FORMALLY UNDECIDABLE PROPOSITIONS OF PRINCIPIA MATHEMATICA AND RELATED SYSTEMS

WINNER of the PULITZER PRIZE

GÖDEL, ESCHER, BACH:
an Eternal Golden Braid
DOUGLAS R. HOFSTADTER
A metaphorical fugue on minds and machines in the spirit of Lewis Carroll

Stanford Encyclopedia of Philosophy
■ Browse ■ About ■ Support SEP

Gödel's Incompleteness Theorems
First published Mon Nov 11, 2013; substantive revision Tue Jan 20, 2015

CAMBRIDGE INTRODUCTIONS TO PHILOSOPHY
An Introduction to Gödel's Theorems
PETER SMITH
SECOND EDITION
CAMBRIDGE

Herbert B. Enderton
A MATHEMATICAL INTRODUCTION TO LOGIC
SECOND EDITION

Studies in Logic
Incompleteness in the Land of Sets
Melvin Fitting

Gödel's Incompleteness Theorems
RAYMOND M. SMULLYAN

Perspectives in Mathematical Logic
Petr Hájek
Pavel Pudlák
Metamathematics of First-Order Arithmetic
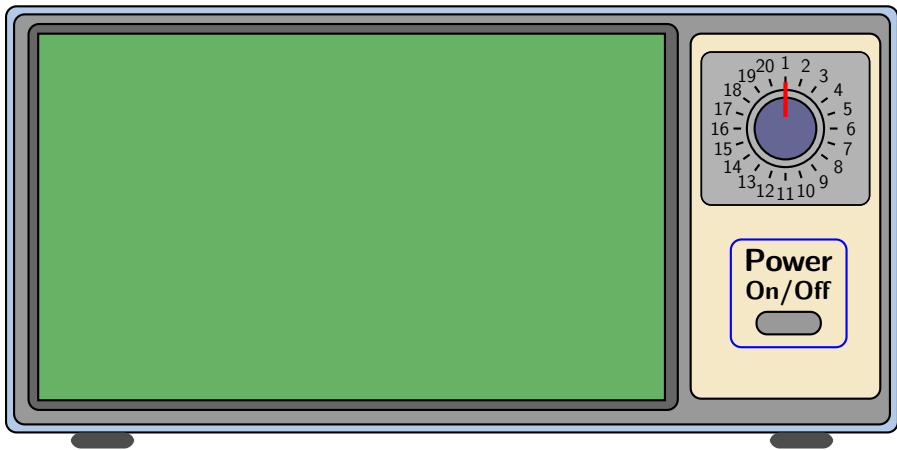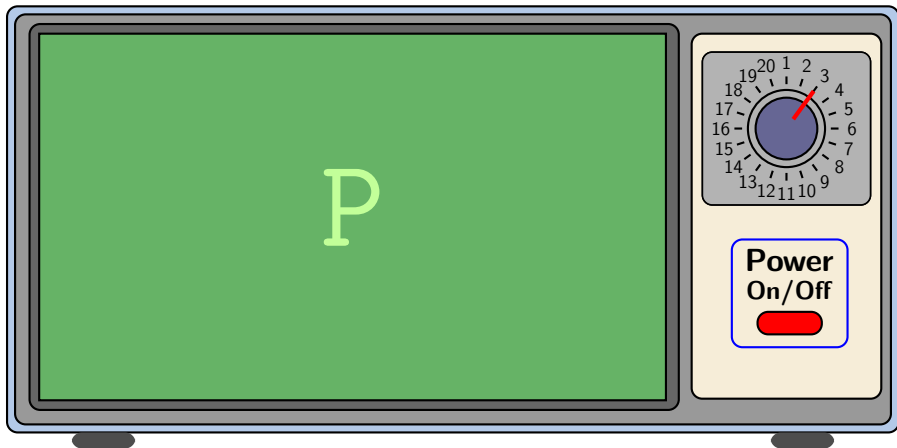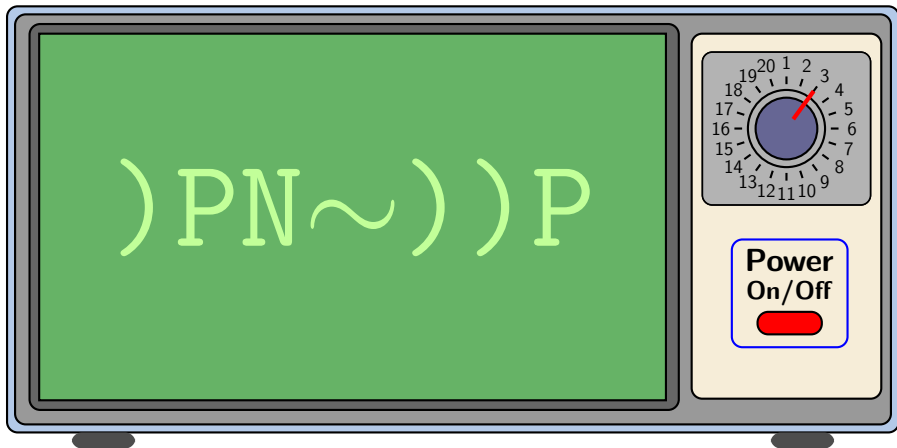Springer-Verlag Berlin Heidelberg GmbH

Smullyan's machine

R. Smullyan. *Chapter 1: The General Idea Behind Gödel's Proof, In* Gödel's Incompleteness Theorems. Oxford University Press, 1992.

Consider a machine that displays strings of the following symbols:

$$) \qquad (\qquad P \qquad N \qquad \sim$$
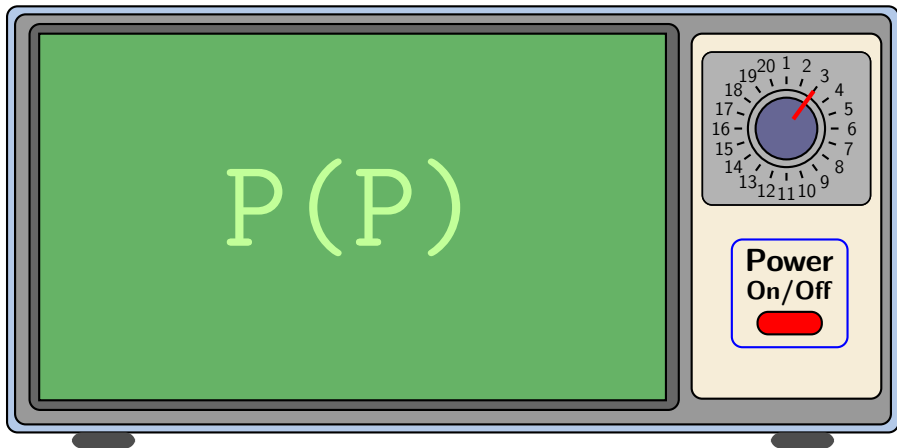
An **expression** is any finite string of) , (, P , N or $\sim$.

An **expression** is any finite string of), (, P, N or $\sim$.

Given an expression X, the **norm** of X is $X(X)$.

An **expression** is any finite string of) , (, P , N or $\sim$.

Given an expression X, the **norm** of X is $X(X)$.

Question: What are the norms of $\sim$P, N)P, P(P), PN and $\sim$PN?

An **expression** is any finite string of), (, P, N or $\sim$.

Given an expression X, the **norm** of X is $X(X)$.

Question: What are the norms of $\sim$P, N)P, P(P), PN and $\sim$PN?

Answer:

1. The norm of $\sim$P is $\sim$P($\sim$P)
2. The norm of N)P is N)P(N)P)
3. The norm of P(P) is P(P)(P(P))

An **expression** is any finite string of), (, P, N or $\sim$.

Given an expression X, the **norm** of X is $X(X)$.

Question: What are the norms of $\sim$P, N)P, P(P), PN and $\sim$PN?

Answer:

1. The norm of $\sim$P is $\sim$P($\sim$P)
2. The norm of N)P is N)P(N)P)
3. The norm of P(P) is P(P)(P(P))
4. The norm of PN is PN(PN)
5. The norm of $\sim$PN is $\sim$PN($\sim$PN)

A **statement** is any expression of the following form:

```
P(X)
~P(X)
PN(X)
 ~PN(X)
```

Statement   is true if...

P(X)

~P(X)

PN(X)

~PN(X)

Statement	is true if...
P(X)	the expression X is printable.
~P(X)	the expression X is not printable.
PN(X)	
~PN(X)	

Statement	is true if…
P(X)	the expression X is printable.
~P(X)	the expression X is not printable.
PN(X)	the norm of X is printable.
~PN(X)	the norm of X is not printable.

)PN∼))P
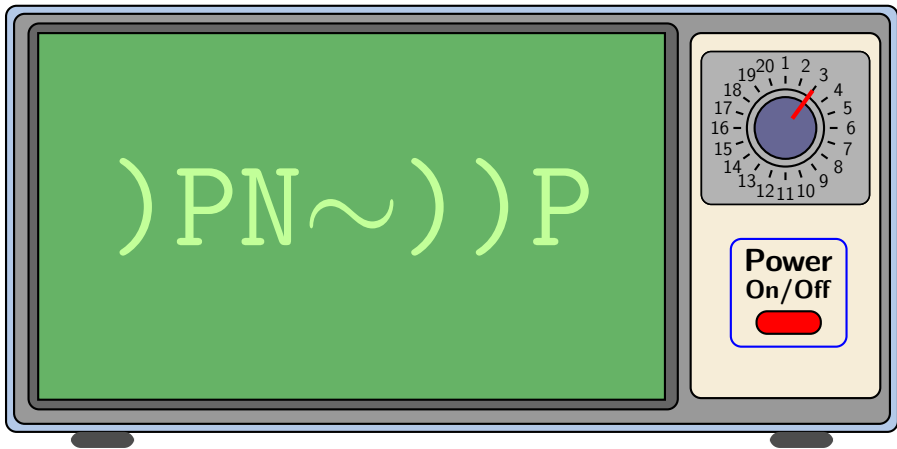
Not a statement, so neither true nor false.

Not a statement, so neither true nor false.

~P(P)

This is false.

$$P(\sim P(P))$$
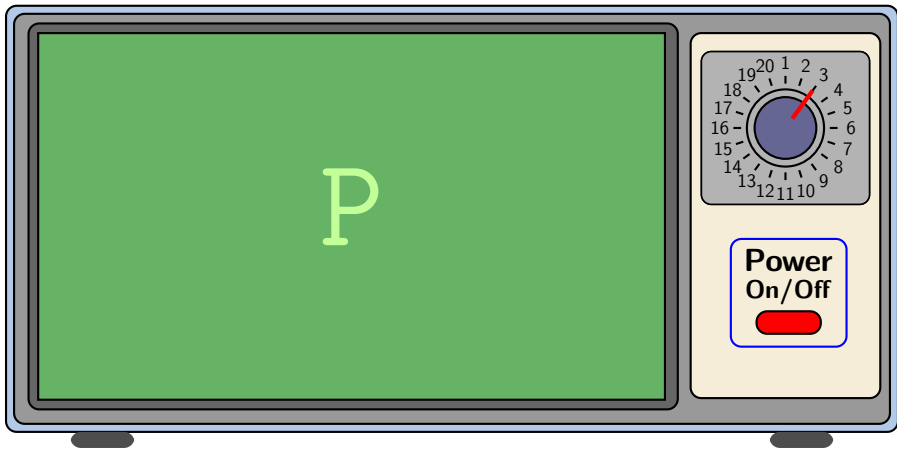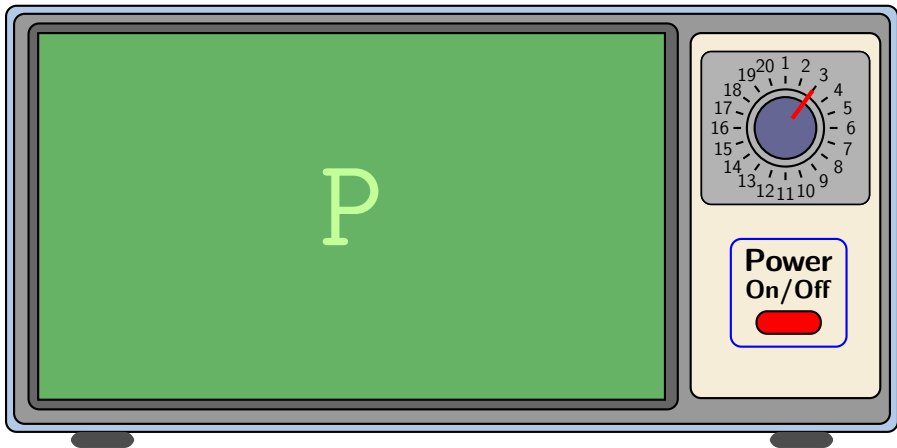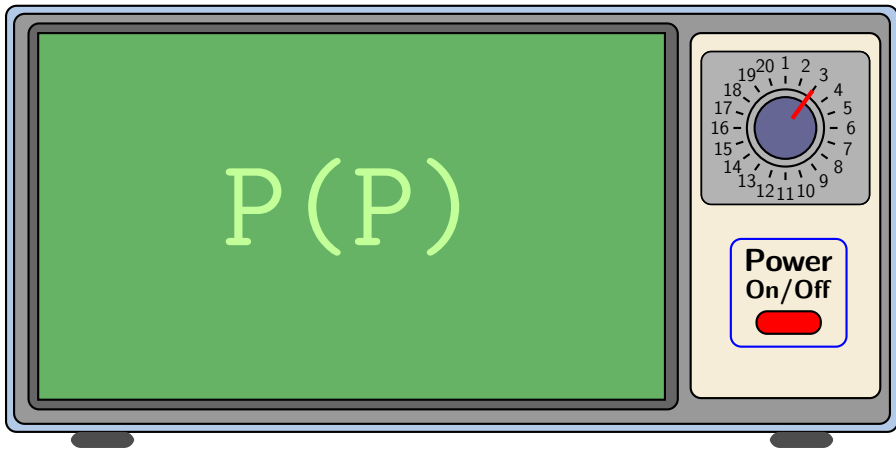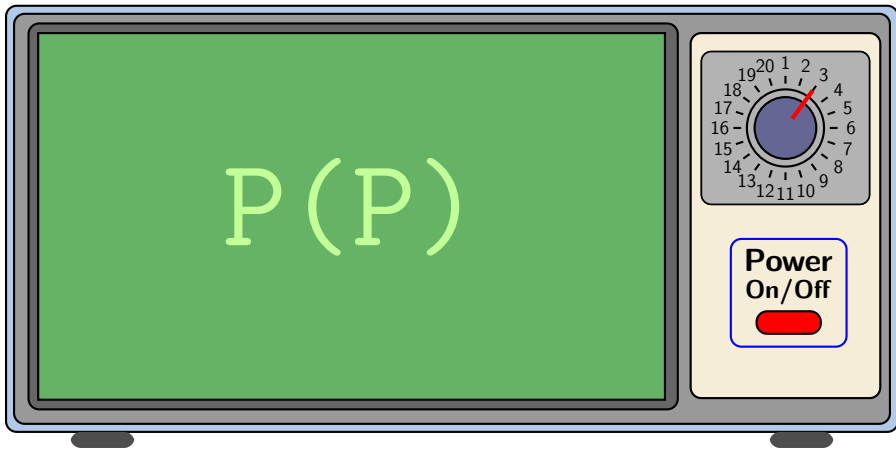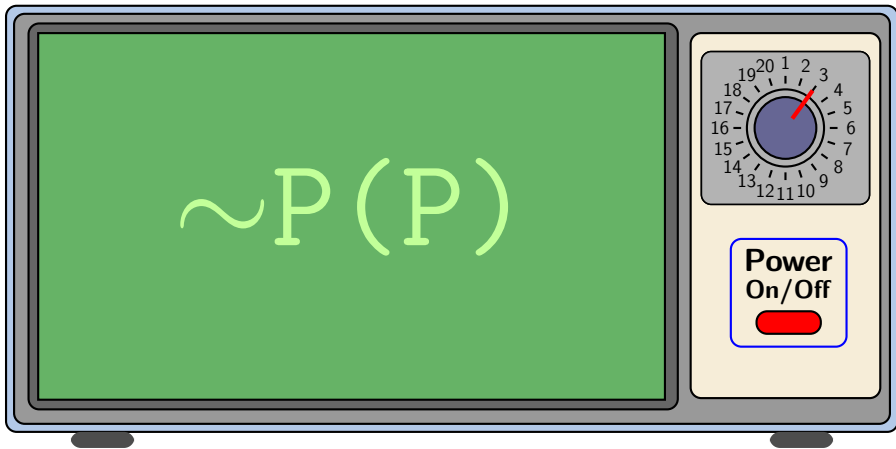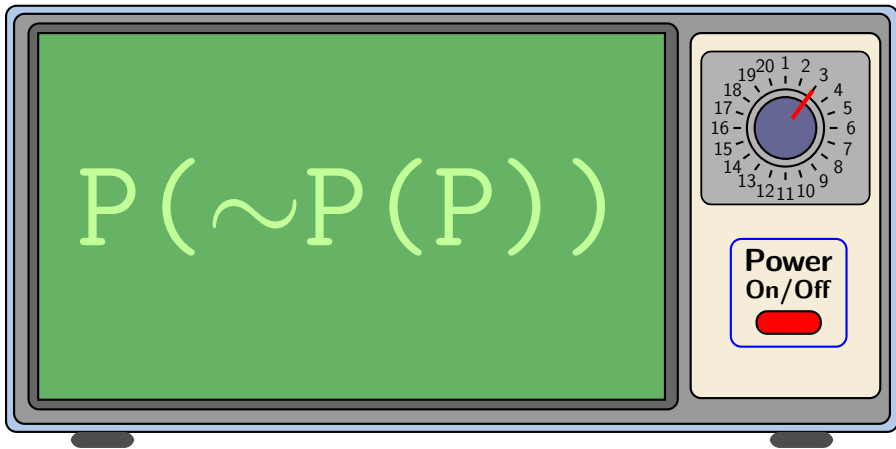
This is true.

Assumption: The machine only prints true statements
(if the machine prints a statement, then it is true).

Assumption: The machine only prints true statements
(if the machine prints a statement, then it is true).

Is it possible to construct a machine that print *all* true statements?
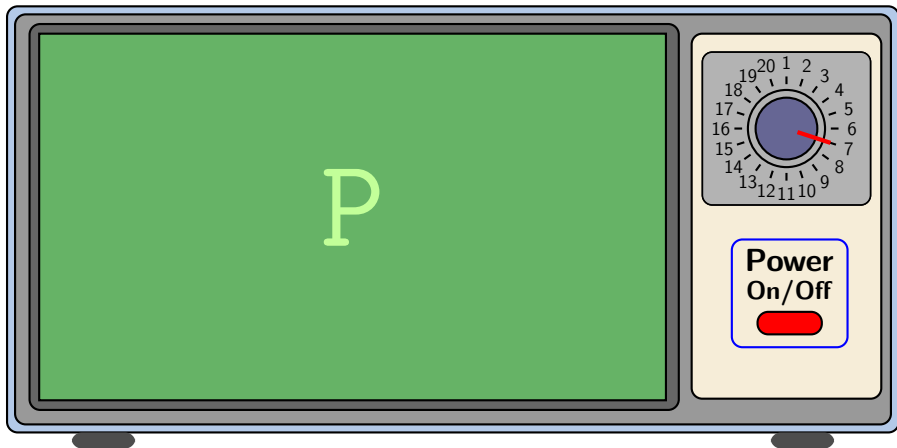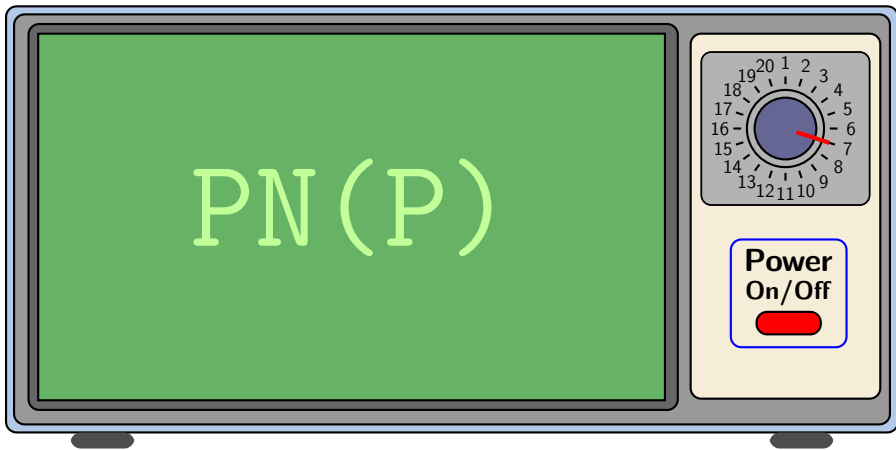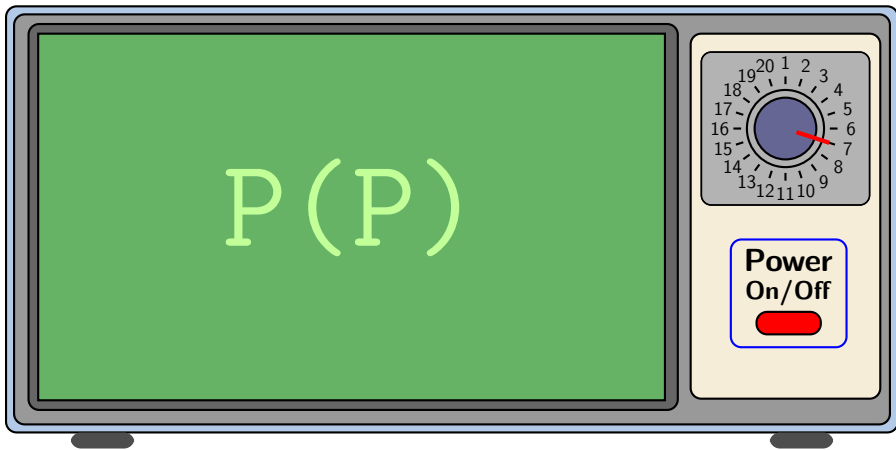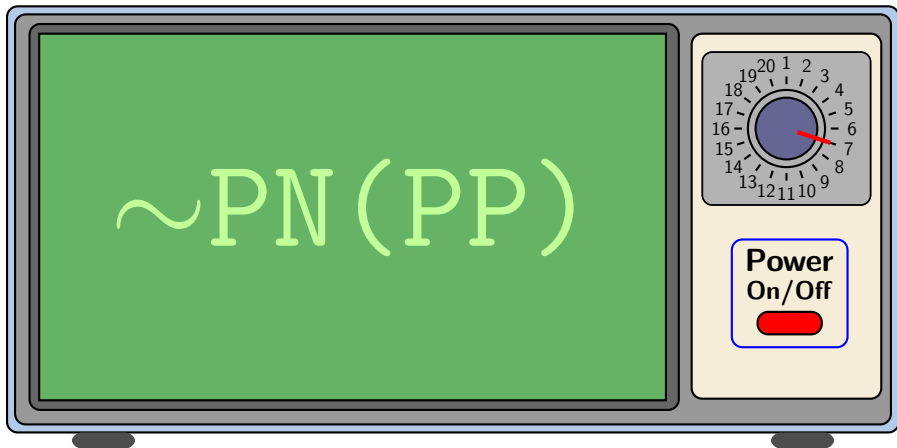
The machine is designed so that PP(PP) will not be printed
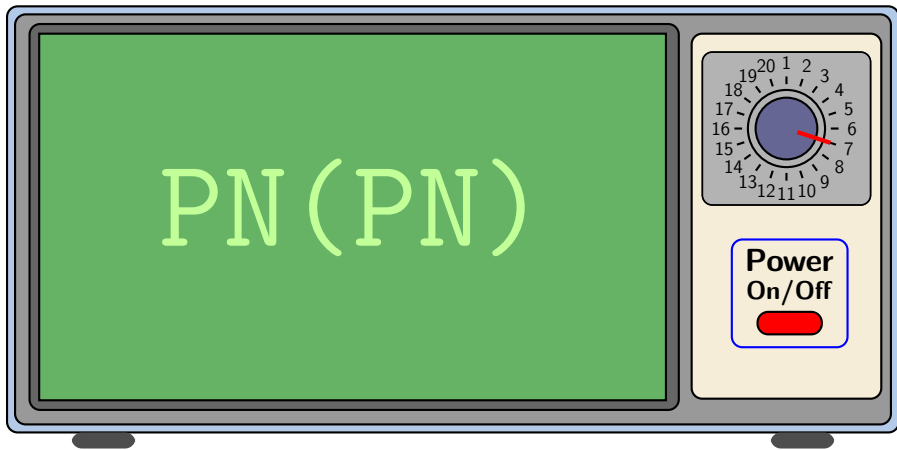
PN(PN) is true

if, and only if,

the norm of PN is printable

if, and only if,

PN(PN) is printable.
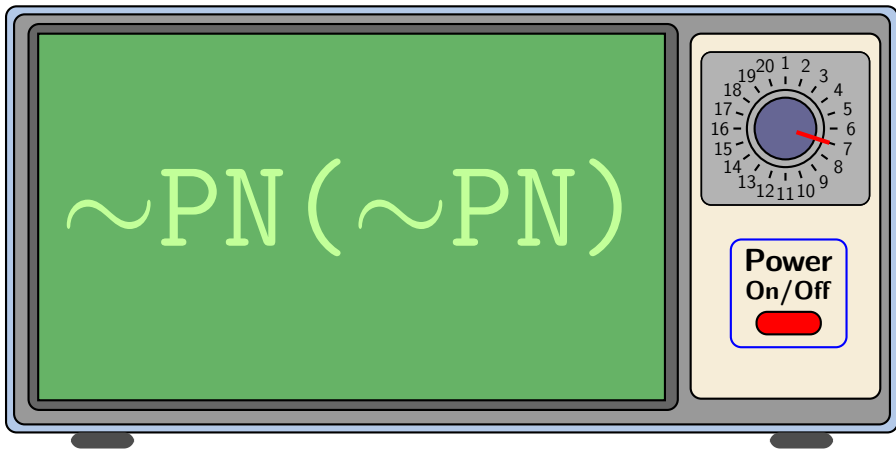
$\sim$PN($\sim$PN) is true

if, and only if,

the norm of $\sim$PN is not printable

if, and only if,

$\sim$PN($\sim$PN) is not printable.

$\sim$PN($\sim$PN) is true if, and only if, $\sim$PN($\sim$PN) is not printable.

$\sim$PN($\sim$PN) is true if, and only if, $\sim$PN($\sim$PN) is not printable.

Two possibilities:

1. The machine is designed to print $\sim$PN($\sim$PN)

2. The machine is designed to not print $\sim$PN($\sim$PN)

~PN(~PN) is true if, and only if, ~PN(~PN) is not printable.

Two possibilities:

1. The machine is designed to print ~PN(~PN): There is a statement that is printable, but not true. (Contradicts the assumption.)

2. The machine is designed to not print ~PN(~PN): There is a statement that is true, but is not printable.
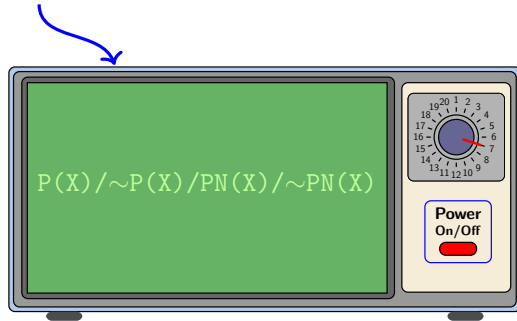
P(X)/∼P(X)/PN(X)/∼PN(X)

Formal system producing statements
about the natural numbers

Formal system producing statements about the natural numbers

Statements about the natural numbers

P(X)/∼P(X)/PN(X)/∼PN(X)

Formal system producing statements about the natural numbers

Statements about the natural numbers

$P(X)/\sim P(X)/PN(X)/\sim PN(X)$

Power On/Off
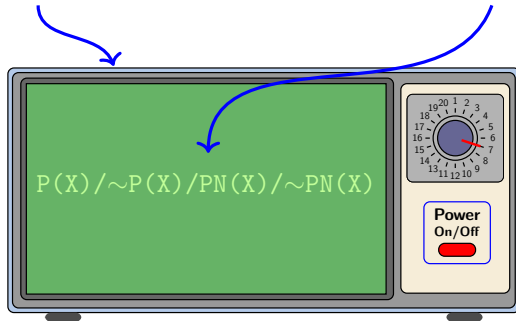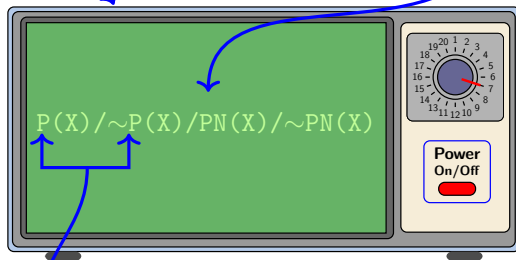
It is/is not provable in the formal system that...

Formal system producing statements about the natural numbers

Statements about the natural numbers

P(X)/∼P(X)/PN(X)/∼PN(X)
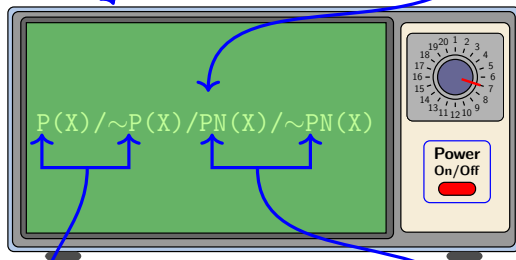
It is/is not provable in the formal system that…

The *diagonal* of the expression is/is not provable…

Background

"…It would seem reasonable, therefore, to surmise that these **axioms and rules of inference are sufficient to decide all mathematical questions which can be formulated in the system concerned**.

"...It would seem reasonable, therefore, to surmise that these **axioms and rules of inference are sufficient to decide all mathematical questions which can be formulated in the system concerned**. In what follows it will be shown that this is not the case, but rather that, in both cited systems, there exists relatively simple problems of the theory of ordinary whole numbers which cannot be decided on the basis of the axioms." (Gödel)

Hilbert's Program

Hilbert's Program had two goals:

1. A complete axiomatization of mathematics, one which will settle every question in mathematics.
2. A proof using strictly finitary means to analyze the formal aspects of the above theory that the axiomatization is *reliable* (i.e., consistent).

R. Zach. *Hilbert's Program*. Stanford Encyclopedia of Philosophy, 2019, https://plato.stanford.edu/entries/hilbert-program/.

- $0 \neq 1$

- $2 \times (1 + 4) = 5 + 3 \times 1 + 2 \times 1 + 0$

- $3 \times 2 = 2 \times 3$

▶ for all n, if $n \neq 0$, then there is a m such that $m + 1 = n$

▶ for all n, m, $n \times m = m \times n$

▶ There is no number smaller than 0

- ▶ There is no biggest prime number

- ▶ there are no a, b, c such that $a^n + b^n = c^n$ for $n > 2$

- every even number is the sum of two prime numbers

- there are infinitely many primes that differ by 2

- for every number n there is a sequence of numbers $k_0, k_1, \ldots, k_m$ such that $k_0 = n$, for each $0 < i \le m$, $k_m = k_{m-1}/2$ if k_{m-1} is even and $k_m = 3k_{m-1} + 1$ if k_{m-1} is odd, and $k_m = 1$
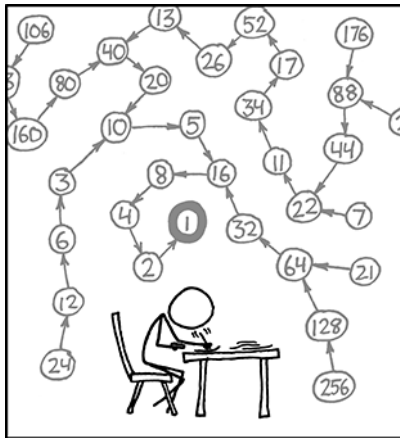
THE COLLATZ CONJECTURE STATES THAT IF YOU PICK A NUMBER, AND IF IT'S EVEN DIVIDE IT BY TWO AND IF IT'S ODD MULTIPLY IT BY THREE AND ADD ONE, AND YOU REPEAT THIS PROCEDURE LONG ENOUGH, EVENTUALLY YOUR FRIENDS WILL STOP CALLING TO SEE IF YOU WANT TO HANG OUT.

Language of Arithmetic $\mathcal{L}_A$

Each of these statements can be expressed in the **language of arithmetic**.

Language of Arithmetic $\mathcal{L}_A$

Each of these statements can be expressed in the **language of arithmetic**.

$$
\begin{array}{ll}
\text{Terms} & 0 \mid x \mid \mathsf{S}(x) \mid (x + y) \mid x \times y \\
\text{Formulas of } \mathcal{L}_A & (t = s) \mid (t < s) \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\forall x)\varphi
\end{array}
$$

Language of Arithmetic $\mathcal{L}_A$

Each of these statements can be expressed in the **language of arithmetic**.

$$
\begin{array}{rl}
\text{Terms} & 0 \mid x \mid \mathsf{S}(x) \mid (x + y) \mid x \times y \\
\text{Formulas of } \mathcal{L}_A & (t = s) \mid (t < s) \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\forall x)\varphi
\end{array}
$$

If we could specify some axioms and inference rules that pin down the number sequence and characterize S, $+$ and $\times$, then we should be able to *decide* any statement about the natural numbers.

The Standard Model

$$\mathcal{N} = (\mathbb{N}, 0, \mathsf{S}, +, *, <)$$

- $0^{\mathcal{N}} = 0$
- $\mathsf{S}^{\mathcal{N}} : \mathbb{N} \to \mathbb{N}$ is the successor function: for all $n \in \mathbb{N}$, $\mathsf{S}^{\mathcal{N}}(n) = n + 1$
- $+^{\mathcal{N}} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is addition: for all $n, m \in \mathbb{N}$, $+^{\mathcal{N}}(n, m) = n + m$
- $\times^{\mathcal{N}} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is multiplication: for all $n, m \in \mathbb{N}$, $\times^{\mathcal{N}}(n, m) = n * m$
- $<^{\mathcal{N}} \subseteq \mathbb{N} \times \mathbb{N}$ is less-than: for all $n, m \in \mathbb{N}$, $(n, m) \in <^{\mathcal{N}}(n, m)$ provided that $n < m$.

Numerals

For each $n \in \mathbb{N}$ we write $\overline{n}$ for the term representing n:

$$\overline{n} \ \text{ is } \ \underbrace{S(\cdots(S(0))\cdots)}_{n \text{ times}}$$

For instance, $\overline{3}$ is $S(S(S(0)))$

To simplify the notation, we often drop the parentheses in the terms $\overline{n}$. For instance, we write $SSS(0)$ instead of $S(S(S(0)))$.

Robinson's **Q**

S1. $\forall x(0 \neq S(x))$

S2. $\forall x \forall y(S(x) = S(y) \rightarrow x = y)$

S3. $\forall x(x \neq 0 \rightarrow \exists y(x = S(y)))$

Robinson's **Q**

S1. $\forall x(0 \neq S(x))$

S2. $\forall x \forall y(S(x) = S(y) \to x = y)$

S3. $\forall x(x \neq 0 \to \exists y(x = S(y)))$

A1. $\forall x(x + 0 = x)$

A2. $\forall x \forall y(x + S(y) = S(x + y))$

Robinson's **Q**

S1. $\forall x (0 \neq \mathsf{S}(x))$

S2. $\forall x \forall y (\mathsf{S}(x) = \mathsf{S}(y) \rightarrow x = y)$

S3. $\forall x (x \neq 0 \rightarrow \exists y (x = \mathsf{S}(y)))$

A1. $\forall x (x + 0 = x)$

A2. $\forall x \forall y (x + \mathsf{S}(y) = \mathsf{S}(x + y))$

M1. $\forall x (x \times 0 = 0)$

M2. $\forall x \forall y (x \times \mathsf{S}(y) = x \times y + x)$

We write $\mathbf{Q} \vdash A$ when there is a derivation of A in which the only open assumptions are the axioms of $\mathbf{Q}$.

Defining $<$

$$x < y \leftrightarrow \exists z(x + \mathsf{S}(z) = y)$$

Exercises

- $\mathbf{Q} \vdash \overline{1} \neq \overline{2}$
- $\mathbf{Q} \vdash \overline{1} + \overline{1} = \overline{2}$
- $\mathbf{Q} \vdash 0 + \overline{3} = \overline{3} + 0$

Exercises

- $\mathbf{Q} \vdash \overline{1} \neq \overline{2}$
- $\mathbf{Q} \vdash \overline{1} + \overline{1} = \overline{2}$
- $\mathbf{Q} \vdash 0 + \overline{3} = \overline{3} + 0$

- For all closed terms s, t,
    - if $\mathcal{N} \models s = t$, then $\mathbf{Q} \vdash s = t$
    - if $\mathcal{N} \models s \neq t$, then $\mathbf{Q} \vdash s \neq t$

Exercises

- $\mathbf{Q} \vdash \overline{1} \neq \overline{2}$
- $\mathbf{Q} \vdash \overline{1} + \overline{1} = \overline{2}$
- $\mathbf{Q} \vdash 0 + \overline{3} = \overline{3} + 0$

- For all closed terms s, t,
    - if $\mathcal{N} \models s = t$, then $\mathbf{Q} \vdash s = t$
    - if $\mathcal{N} \models s \neq t$, then $\mathbf{Q} \vdash s \neq t$

- $\mathbf{Q} \vdash \forall x(x + 0 = x)$

Exercises

- $\mathbf{Q} \vdash \overline{1} \neq \overline{2}$
- $\mathbf{Q} \vdash \overline{1} + \overline{1} = \overline{2}$
- $\mathbf{Q} \vdash 0 + \overline{3} = \overline{3} + 0$

- For all closed terms s, t,
  - if $\mathcal{N} \models s = t$, then $\mathbf{Q} \vdash s = t$
  - if $\mathcal{N} \models s \neq t$, then $\mathbf{Q} \vdash s \neq t$

- $\mathbf{Q} \vdash \forall x (x + 0 = x)$

- For all $n \in \mathbb{N}$, $\mathbf{Q} \vdash 0 + \overline{n} = \overline{n}$

Exercises

- $\mathbf{Q} \vdash \overline{1} \neq \overline{2}$
- $\mathbf{Q} \vdash \overline{1} + \overline{1} = \overline{2}$
- $\mathbf{Q} \vdash 0 + \overline{3} = \overline{3} + 0$

- For all closed terms s, t,
    - if $\mathcal{N} \models s = t$, then $\mathbf{Q} \vdash s = t$
    - if $\mathcal{N} \models s \neq t$, then $\mathbf{Q} \vdash s \neq t$

- $\mathbf{Q} \vdash \forall x(x + 0 = x)$

- For all $n \in \mathbb{N}$, $\mathbf{Q} \vdash 0 + \overline{n} = \overline{n}$

- $\mathbf{Q} \not\vdash \forall x(0 + x = x)$

Peano Arithmetic (**PA**)

The axioms of **PA** (Peano Arithmetic) are all the axioms of **Q** with every instance of the following axiom schema:

Induction Scheme: For all formulas φ of $\mathcal{L}_A$,

$$(\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(S(x)))) \to \forall x\, \varphi(x)$$

We write **PA** $\vdash \varphi$ when there is a derivation of φ in which the only open assumptions are the axioms of **PA**.

Exercises, continued

- **PA** $\vdash \forall x(0 + x = x)$

The Theory of True Arithmetic

True arithmetic: $Th(\mathcal{N}) = \{\varphi \mid \mathcal{N} \models \varphi\}$

1. Is there a computational procedure we can use to test if a sentence is in $Th(\mathcal{N})$?

The Theory of True Arithmetic

True arithmetic: $Th(\mathcal{N}) = \{\varphi \mid \mathcal{N} \models \varphi\}$

1. Is there a computational procedure we can use to test if a sentence is in $Th(\mathcal{N})$?

A **theory** is a set of sentences that is closed under entailment, i.e., **T** is a theory if $\mathbf{T} = \{\varphi \mid \mathbf{T} \models \varphi\}$

A theory is **axiomatizable** if there is a *decidable* set of sentences $\mathbf{T}_0$ such that $\mathbf{T} = \{\varphi \mid \mathbf{T}_0 \models \varphi\}$

The Theory of True Arithmetic

True arithmetic: $Th(\mathcal{N}) = \{\varphi \mid \mathcal{N} \models \varphi\}$

1. Is there a computational procedure we can use to test if a sentence is in $Th(\mathcal{N})$?

A **theory** is a set of sentences that is closed under entailment, i.e., **T** is a theory if $\mathbf{T} = \{\varphi \mid \mathbf{T} \models \varphi\}$

A theory is **axiomatizable** if there is a *decidable* set of sentences $\mathbf{T}_0$ such that $\mathbf{T} = \{\varphi \mid \mathbf{T}_0 \models \varphi\}$

2. Is there an axiomatizable theory **T** such that $\mathbf{T} = Th(\mathcal{N})$? This is equivalent to asking whether **T** is **complete**: For every sentence φ, either $\mathbf{T} \models \varphi$ or $\mathbf{T} \models \neg\varphi$.

The answer to both questions is no.

The answer to both questions is no.

Gödel's first incompleteness theorem (informal statement):
Any consistent formal theory within which a certain amount of elementary
arithmetic can be carried out is **incomplete**.

Arithmetic Hierarchy

▶ A quantifier is **bounded** if it is the form '$\forall x \leq t$' or '$\exists x \leq t$', where t is a term not involving x.

▶ A formula is a **bounded formula** (denoted Δ_0^0) if all of its quantifiers are bounded.

Arithmetic Hierarchy

- A quantifier is **bounded** if it is the form '$\forall x \leq t$' or '$\exists x \leq t$', where t is a term not involving x.

- A formula is a **bounded formula** (denoted Δ_0^0) if all of its quantifiers are bounded.

- For $n \geq 0$, the classes of formulas Σ_n^0 and Π_n^0 are defined as follows:
  - $\Sigma_0^0 = \Pi_0^0 = \Delta_0^0$.
  - Σ_{n+1}^0 is the set of formulas of the form $\exists \vec{x} \varphi$ where φ is a Π_n^0 formula and $\vec{x}$ is a (possibly empty) list of variables.
  - Π_{n+1}^0 is the set of formulas of the form $\forall \vec{x} \varphi$ where φ is a Σ_n^0 formula and $\vec{x}$ is a (possibly empty) list of variables.

Definition

Σ_1^0-sound A theory **T** is Σ_1^0-**sound** iff for every Σ_1^0-formula φ, if **T** $\vdash \varphi$, then φ is true (in the standard model).

Definition

Σ_1^0-complete A theory **T** is Σ_1^0-**complete** iff for every Σ_1^0-formula φ, if φ is true (in the standard model), then **T** $\vdash \varphi$.

Proposition

PA (in fact, even **Q**) is Σ_1^0-complete.

Theorem (Gödel's First Incompleteness Theorem)

Assume that **PA** is Σ_1^0-sound. Then there is a Π_1^0-sentence φ such that **PA** neither proves φ nor $\neg\varphi$.

Theorem (Gödel's Second Incompleteness Theorem)

Assume that **PA** is consistent. Then **PA** cannot prove $\mathrm{Con}_{\textbf{PA}}$.

$\mathrm{Con}_{\textbf{PA}}$ *is a Π_1^0-statement that informally asserts "for all x, x does not code a proof of a contradiction from the axioms of* **PA**"

- ▶ Gödel numbering
- ▶ Gödel-Carnap Fixed Point Theorem
- ▶ (Naming systems)
- ▶ Representing functions/relations

Gödel Numbering

Gödel-numbering assigns numbers to the syntactic objects of a logic (i.e., the terms, the formulas, and the derivations).

Suppose that χ is a syntactic object (i.e., a term, formula or a derivation). We use the following notation:

$gn(\chi)$: The Gödel number of χ (an integer)

$\ulcorner \chi \urcorner$: The numeral of the Gödel number of χ (a numeral). That is:

$$\ulcorner \chi \urcorner \equiv \overline{gn(\chi)}$$

Fixed-Point Theorem

Theorem (Gödel-Carnap Fixed-Point Theorem)

Let $A(x)$ be any formula of $\mathcal{L}_A$ with one free variable x. Then there is a sentence B such that

$$\mathbf{Q} \vdash B \leftrightarrow A(\ulcorner B \urcorner).$$

Substitution

Suppose that A is a formula where x is a free variable. We write $A(x)$ to when the formula A has at most one free variable x.

If t is a term, then $A(x)[x/t]$ is A with every instance of x replaced with t. We sometimes abuse notation and write $A(t)$ instead of $A(x)[x/t]$.

Substitution

Suppose that A is a formula where x is a free variable. We write $A(x)$ to when the formula A has at most one free variable x.

If t is a term, then $A(x)[x/t]$ is A with every instance of x replaced with t. We sometimes abuse notation and write $A(t)$ instead of $A(x)[x/t]$.

Let $Sub : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a function, where for each $n, m \in \mathbb{N}$, $Sub(n, m)$ is the code of $\alpha(x)[x/\overline{m}]$ where n is the code of $\alpha(x)$. So, for any formula $\alpha(x)$ and $m \in \mathbb{N}$:

$$Sub(gn(\alpha(x)), m) = gn(\alpha(\overline{m}))$$

We sketch a proof under the assumption that sub is a function symbol in the language $\mathcal{L}_A$ and the theory **Q** "represents" *Sub* in the following sense:

For any formula $A(x)$ and $n \in \mathbb{N}$,

$$\mathbf{Q} \vdash \text{sub}(\ulcorner A(x) \urcorner, \overline{n}) = \ulcorner A(\overline{n}) \urcorner$$

- Let $A^*(x)$ be $A(\text{sub}(x, x))$
  Let B be $A^*(\ulcorner A^*(x) \urcorner)$

- $A^*(\ulcorner A^*(x) \urcorner)$ is the formula $A(\text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner))$

- Let $A^*(x)$ be $A(\mathrm{sub}(x, x))$
  Let B be $A^*(\ulcorner A^*(x) \urcorner)$

- $A^*(\ulcorner A^*(x) \urcorner)$ is the formula $A(\mathrm{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner))$

$$\mathbf{Q} \vdash \qquad \mathrm{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner) \qquad = \qquad \ulcorner A^*(\ulcorner A^*(x) \urcorner) \urcorner$$

- Let $A^*(x)$ be $A(\text{sub}(x, x))$

  Let B be $A^*(\ulcorner A^*(x) \urcorner)$

- $A^*(\ulcorner A^*(x) \urcorner)$ is the formula $A(\text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner))$

$\mathbf{Q} \vdash \qquad \text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner) \qquad = \qquad \ulcorner A^*(\ulcorner A^*(x) \urcorner) \urcorner$

$\mathbf{Q} \vdash \quad A(\, \text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner) \,) \quad \leftrightarrow \quad A(\, \ulcorner A^*(\ulcorner A^*(x) \urcorner) \urcorner \,)$

- Let $A^*(x)$ be $A(\text{sub}(x, x))$
  Let B be $A^*(\ulcorner A^*(x) \urcorner)$

- $A^*(\ulcorner A^*(x) \urcorner)$ is the formula $A(\text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner))$

$$\mathbf{Q} \vdash \quad \text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner) \quad = \quad \ulcorner A^*(\ulcorner A^*(x) \urcorner) \urcorner$$

$$\mathbf{Q} \vdash \quad A(\text{sub}(\ulcorner A^*(x) \urcorner, \ulcorner A^*(x) \urcorner)) \quad \leftrightarrow \quad A(\ulcorner A^*(\ulcorner A^*(x) \urcorner) \urcorner)$$

$$\mathbf{Q} \vdash \qquad\qquad\qquad\qquad\qquad B \quad \leftrightarrow \quad A(\ulcorner B \urcorner)$$