Interpretability and the arithmetized completeness theorem

倉橋　太志 (Taishi Kurahashi)

神戸大学 (Kobe University)

2011 年 12 月 09–11 日　数学基礎論若手の会

- *Interpretability* **is used to prove relative consitency, decidability and undecidability of theories.**

- *Interpretability* **is used to prove relative consitency, decidability and undecidability of theories.**
- **The notion of interpretability was explicitly introduced by Tarski(1954) and systematically investigated by Feferman and Orey(1960).**

- *Interpretability* **is used to prove relative consitency, decidability and undecidability of theories.**
- **The notion of interpretability was explicitly introduced by Tarski(1954) and systematically investigated by Feferman and Orey(1960).**
- **In this talk, we introduce a result about interpretability proved by using the arithmetized completeness theorem in Feferman(1960).**

Contents

Definition

$\mathcal{L}, \mathcal{L}'$: languages, t: $\mathsf{Fml}_{\mathcal{L}} \to \mathsf{Fml}_{\mathcal{L}'}$.

t is a **translation** of $\mathcal{L}$ into $\mathcal{L}'$
$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

Definition

$\mathcal{L}, \mathcal{L}'$: languages, t: $\mathsf{Fml}_{\mathcal{L}} \to \mathsf{Fml}_{\mathcal{L}'}$.

t is a **translation** of $\mathcal{L}$ into $\mathcal{L}'$

$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

- $t(x = y) \equiv x = y$;
- $\forall c \in \mathcal{L}$: constant, $\exists \eta_c(x)$: $\mathcal{L}'$-formula s.t. $t(x = c) \equiv \eta_c(x)$;
- $\cdots$;

Definition

$\mathcal{L}, \mathcal{L}'$: languages, t: $\mathsf{Fml}_{\mathcal{L}} \to \mathsf{Fml}_{\mathcal{L}'}$.

t is a **translation** of $\mathcal{L}$ into $\mathcal{L}'$

$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

- $t(x = y) \equiv x = y$;
- $\forall c \in \mathcal{L}$: constant, $\exists \eta_c(x)$: $\mathcal{L}'$-formula s.t. $t(x = c) \equiv \eta_c(x)$;
- $\cdots$;
- $t(\neg\varphi) \equiv \neg t(\varphi)$ for any $\mathcal{L}$-formula φ;
- $t(\varphi \vee \psi) \equiv t(\varphi) \vee t(\psi)$ for any $\mathcal{L}$-formulas φ, ψ;
- $\cdots$;

Definition

$\mathcal{L}, \mathcal{L}'$: languages, t: $\mathsf{Fml}_{\mathcal{L}} \to \mathsf{Fml}_{\mathcal{L}'}$.
t is a **translation** of $\mathcal{L}$ into $\mathcal{L}'$
$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

- $t(x = y) \equiv x = y$;
- $\forall c \in \mathcal{L}$: constant, $\exists \eta_c(x)$: $\mathcal{L}'$-formula s.t. $t(x = c) \equiv \eta_c(x)$;
- $\cdots$;
- $t(\neg\varphi) \equiv \neg t(\varphi)$ for any $\mathcal{L}$-formula φ;
- $t(\varphi \vee \psi) \equiv t(\varphi) \vee t(\psi)$ for any $\mathcal{L}$-formulas φ, ψ;
- $\cdots$;
- $\exists d(x)$: $\mathcal{L}'$-formula s.t.
  $t(\exists x \varphi(x)) \equiv \exists x(d(x) \wedge t(\varphi(x)))$ for any $\mathcal{L}$-formula $\varphi(x)$.

$\mathcal{L}_T :=$ the language of T.

$\mathcal{L}_T$:= the language of T.

Definition

S, T: theories, t: translation of $\mathcal{L}_S$ into $\mathcal{L}_T$.

t is an **interpretation** of S in T
$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

1. $T \vdash \exists x d(x)$;
2. $\forall \varphi$: $\mathcal{L}_S$-formula, $(S \vdash \varphi \Rightarrow T \vdash t(\varphi))$.

$\mathcal{L}_T$:= the language of T.

Definition

S, T: theories, t: translation of $\mathcal{L}_S$ into $\mathcal{L}_T$.
t is an **interpretation** of S in T
$\overset{\text{def.}}{\Leftrightarrow}$ t satisfies the following conditions:

1. $T \vdash \exists x d(x)$;
2. $\forall \varphi$: $\mathcal{L}_S$-formula, $(S \vdash \varphi \Rightarrow T \vdash t(\varphi))$.

Definition

S is **interpretable** in T ($S \leq T$)
$\overset{\text{def.}}{\Leftrightarrow}$ $\exists t$: interpretation of S in T.

It is easy to check the following propositions.

- **If $S \leq T$ and T is consistent, then S is consistent.**
- **If $S \leq T$ and T is decidable, then S is decidable.**

It is easy to check the following propositions.

- If $S \leq T$ and T is consistent, then S is consistent.
- If $S \leq T$ and T is decidable, then S is decidable.

- $\mathcal{L}_A := \{0, S, +, \times, <\}$: language of arithmetic.
- PA: Peano arithmetic (Basic axioms of arithmetic with induction scheme for all $\mathcal{L}_A$-formulas).
- ZF: Zermelo-Fraenkel set theory (Inf: Axiom of infinity).

It is easy to check the following propositions.

- If $S \leq T$ and T is consistent, then S is consistent.
- If $S \leq T$ and T is decidable, then S is decidable.

- $\mathcal{L}_A := \{0, S, +, \times, <\}$: language of arithmetic.
- PA: Peano arithmetic (Basic axioms of arithmetic with induction scheme for all $\mathcal{L}_A$-formulas).
- ZF: Zermelo-Fraenkel set theory (Inf: Axiom of infinity).

- $\mathsf{ZF} - \mathsf{Inf} \leq \mathsf{PA}$.
- $\mathsf{ZF} \not\leq \mathsf{PA}$.
- $\mathsf{PA} + \mathsf{Con}_{\mathsf{PA}} \not\leq \mathsf{PA}$.
- $\mathsf{PA} + \neg\mathsf{Con}_{\mathsf{PA}} \leq \mathsf{PA}$.
- $\cdots$.

Question

Is there any $\mathcal{L}_A$-sentence φ s.t. ZF $\leq$ PA $+ \varphi$?

Question

Is there any $\mathcal{L}_A$-sentence φ s.t. ZF $\leq$ PA $+ \varphi$?

Answer

Yes!

Question

Is there any $\mathcal{L}_A$-sentence φ s.t. ZF $\leq$ PA $+ \varphi$?

Answer

Yes!

- **In fact,**

$$\mathbf{ZF} \leq \mathbf{PA} + \mathbf{Con_{ZF}}$$

by Feferman's theorem.

Question

Is there any $\mathcal{L}_A$-sentence φ s.t. ZF $\leq$ PA $+ \varphi$?

Answer

Yes!

- **In fact,**

$$\mathsf{ZF} \leq \mathsf{PA} + \mathsf{Con}_{\mathsf{ZF}}$$

by Feferman's theorem.

- **Feferman's theorem is proved by using the arithmetized completeness theorem.**

Countable completeness theorem

T: theory with a countable language $\mathcal{L}$.

If T is consistent, then T has a model.

Countable completeness theorem

T: theory with a countable language $\mathcal{L}$.

If T is consistent, then T has a model.

Proof(outline)

- $C := \{c_n \mid n \in \omega\}$: **set of new constants.**
- $\mathcal{L}_C := \mathcal{L} \cup C$.

Countable completeness theorem

T: theory with a countable language $\mathcal{L}$.

If T is consistent, then T has a model.

Proof(outline)

- $C := \{c_n \mid n \in \omega\}$: **set of new constants.**
- $\mathcal{L}_C := \mathcal{L} \cup C$.
- $\{\varphi_n(x_n)\}_{n\in\omega}$: **primitive recursive(p.r.) enumeration of all $\mathcal{L}_C$-formulas with one free-variable.**
- $\exists Z := \{\exists x_n \varphi_n(x_n) \to \varphi_n(c_{i_n}) \mid n \in \omega\}$: **p.r. set s.t. $T + Z$ is a conservative extension of T.**

Countable completeness theorem

T: theory with a countable language $\mathcal{L}$.

If T is consistent, then T has a model.

Proof(outline)

- $C := \{c_n \mid n \in \omega\}$: **set of new constants.**
- $\mathcal{L}_C := \mathcal{L} \cup C$.
- $\{\varphi_n(x_n)\}_{n\in\omega}$: **primitive recursive(p.r.) enumeration of all $\mathcal{L}_C$-formulas with one free-variable.**
- $\exists Z := \{\exists x_n \varphi_n(x_n) \to \varphi_n(c_{i_n}) \mid n \in \omega\}$: **p.r. set s.t. $T + Z$ is a conservative extension of T.**
- **Then $T + Z$ is consistent (Henkin extension).**

- $\{\theta_n\}_{n\in\omega}$: **p.r. enumeration of all** $\mathcal{L}_C$**-sentences.**

- $\{\theta_n\}_{n\in\omega}$: **p.r. enumeration of all $\mathcal{L}_C$-sentences.**
- $X_0 := T + Z$;

$$X_{n+1} := \begin{cases} X_n \cup \{\theta_n\} & \text{if } X_n \cup \{\theta_n\} \text{ is consistent;} \\ X_n \cup \{\neg\theta_n\} & \text{otherwise.} \end{cases}$$

- $X := \bigcup_{n\in\omega} X_n$. X **is Henkin complete.**

- $\{\theta_n\}_{n\in\omega}$: p.r. enumeration of all $\mathcal{L}_C$-sentences.
- $X_0 := T + Z$;

$$X_{n+1} := \begin{cases} X_n \cup \{\theta_n\} & \text{if } X_n \cup \{\theta_n\} \text{ is consistent;} \\ X_n \cup \{\neg\theta_n\} & \text{otherwise.} \end{cases}$$

- $X := \bigcup_{n\in\omega} X_n$. X is Henkin complete.
- Define an equivalence relation $\sim$ on C by $c \sim d :\Leftrightarrow c = d \in X$.

- $\{\theta_n\}_{n\in\omega}$: p.r. enumeration of all $\mathcal{L}_C$-sentences.
- $X_0 := T + Z$;

$$X_{n+1} := \begin{cases} X_n \cup \{\theta_n\} & \text{if } X_n \cup \{\theta_n\} \text{ is consistent;} \\ X_n \cup \{\neg\theta_n\} & \text{otherwise.} \end{cases}$$

- $X := \bigcup_{n\in\omega} X_n$. X is Henkin complete.
- Define an equivalence relation $\sim$ on C by $c \sim d :\Leftrightarrow c = d \in X$.
- Define a structure $\mathcal{M}$ by $|\mathcal{M}| := C/\sim$ and $R^{\mathcal{M}}([c_0], \ldots, [c_n]) :\Leftrightarrow R(c_0, \ldots, c_n) \in X$ etc...

- $\{\theta_n\}_{n\in\omega}$: p.r. enumeration of all $\mathcal{L}_C$-sentences.
- $X_0 := T + Z$;

$$X_{n+1} := \begin{cases} X_n \cup \{\theta_n\} & \text{if } X_n \cup \{\theta_n\} \text{ is consistent;} \\ X_n \cup \{\neg\theta_n\} & \text{otherwise.} \end{cases}$$

- $X := \bigcup_{n\in\omega} X_n$. X is Henkin complete.
- Define an equivalence relation $\sim$ on C by $c \sim d :\Leftrightarrow c = d \in X$.
- Define a structure $\mathcal{M}$ by $|\mathcal{M}| := C/\sim$ and $R^{\mathcal{M}}([c_0], \dots, [c_n]) :\Leftrightarrow R(c_0, \dots, c_n) \in X$ etc...
- $\forall\varphi$: $\mathcal{L}_C$-sentence, $(\mathcal{M} \models \varphi \Leftrightarrow \varphi \in X)$.
- $\mathcal{M}$ is a model of T.

- $\{\theta_n\}_{n\in\omega}$: p.r. enumeration of all $\mathcal{L}_C$-sentences.
- $X_0 := T + Z$;

$$X_{n+1} := \begin{cases} X_n \cup \{\theta_n\} & \text{if } X_n \cup \{\theta_n\} \text{ is consistent;} \\ X_n \cup \{\neg\theta_n\} & \text{otherwise.} \end{cases}$$

- $X := \bigcup_{n\in\omega} X_n$. X is Henkin complete.
- Define an equivalence relation $\sim$ on C by $c \sim d :\Leftrightarrow c = d \in X$.
- Define a structure $\mathcal{M}$ by $|\mathcal{M}| := C/\sim$ and $R^{\mathcal{M}}([c_0], \ldots, [c_n]) :\Leftrightarrow R(c_0, \ldots, c_n) \in X$ etc...
- $\forall\varphi$: $\mathcal{L}_C$-sentence, $(\mathcal{M} \models \varphi \Leftrightarrow \varphi \in X)$.
- $\mathcal{M}$ is a model of T.

□

First, we arithmetize the notion of the provability.

First, we arithmetize the notion of the provability.

S: **r.e.** $\mathcal{L}$**-theory,** T: $\mathcal{L}_A$**-theory. (**$\mathcal{L}$**: countable)**

First, we arithmetize the notion of the provability.

S: **r.e.** $\mathcal{L}$**-theory,** T: $\mathcal{L}_A$**-theory. (**$\mathcal{L}$: **countable)**

$\exists \sigma(x)$: Σ_1 **formula s.t.** $\forall \varphi$: $\mathcal{L}$**-sentence,**

$$\varphi \in S \Leftrightarrow T \vdash \sigma(\ulcorner \varphi \urcorner).$$

$\sigma(x)$ **is called a numeration of** S **in** T**.**

First, we arithmetize the notion of the provability.

S: r.e. $\mathcal{L}$-theory, T: $\mathcal{L}_A$-theory. ($\mathcal{L}$: countable)

$\exists\sigma(x)$: Σ_1 formula s.t. $\forall\varphi$: $\mathcal{L}$-sentence,

$$\varphi \in S \Leftrightarrow T \vdash \sigma(\ulcorner\varphi\urcorner).$$

$\sigma(x)$ is called a **numeration** of S in T.

For Σ_1 numeration $\sigma(x)$ of S in T,
we can construct a Σ_1 formula $\mathrm{Pr}_\sigma(x)$ s.t. $\forall\varphi$: $\mathcal{L}$-sentence,

$$S \vdash \varphi \Leftrightarrow T \vdash \mathrm{Pr}_\sigma(\ulcorner\varphi\urcorner).$$

$\mathrm{Pr}_\sigma(x)$ is called the **provability predicate** of $\sigma(x)$.

$\sigma(x), \sigma'(x)$: **numerations of S and S' in T respectively.**

Define

- $(\sigma|n)(x) := \sigma(x) \wedge x \leq \bar{n}$.
- $(\sigma \vee \sigma')(x) := \sigma(x) \vee \sigma'(x)$.

$\sigma(x), \sigma'(x)$: numerations of S and S' in T respectively.

Define

- $(\sigma|n)(x) :\equiv \sigma(x) \wedge x \leq \bar{n}$.
- $(\sigma \vee \sigma')(x) :\equiv \sigma(x) \vee \sigma'(x)$.

$(\sigma|n)(x)$ **is a numeration of** $\{\varphi \in S \mid \ulcorner \varphi \urcorner \leq n\}$.

$\sigma(x), \sigma'(x)$: **numerations of** S **and** S' **in** T **respectively.**

Define

- $(\sigma|n)(x) := \sigma(x) \land x \leq \bar{n}$.
- $(\sigma \lor \sigma')(x) := \sigma(x) \lor \sigma'(x)$.

$(\sigma|n)(x)$ **is a numeration of** $\{\varphi \in S \mid \ulcorner \varphi \urcorner \leq n\}$.

Con$_\sigma :\equiv \neg \mathrm{Pr}_\sigma(\ulcorner 0 = \bar{1} \urcorner)$.

$\sigma(x), \sigma'(x)$: **numerations of S and S' in T respectively.**

Define

- $(\sigma|n)(x) :\!= \sigma(x) \land x \leq \bar{n}$.
- $(\sigma \lor \sigma')(x) :\!= \sigma(x) \lor \sigma'(x)$.

$(\sigma|n)(x)$ **is a numeration of** $\{\varphi \in S \mid \ulcorner\varphi\urcorner \leq n\}$.

Con$_\sigma :\equiv \neg\mathrm{Pr}_\sigma(\ulcorner 0 = \bar{1}\urcorner)$.

Theorem

T: **consistent r.e. extension of PA,**
$\sigma(x)$: **numeration of T in T. Then**

$\sigma(x), \sigma'(x)$: numerations of S and S' in T respectively.

Define

- $(\sigma|n)(x) := \sigma(x) \wedge x \leq \bar{n}$.
- $(\sigma \vee \sigma')(x) := \sigma(x) \vee \sigma'(x)$.

$(\sigma|n)(x)$ is a numeration of $\{\varphi \in S \mid \ulcorner \varphi \urcorner \leq n\}$.

$\mathrm{Con}_\sigma :\equiv \neg \mathrm{Pr}_\sigma(\ulcorner 0 = \bar{1} \urcorner)$.

Theorem

T: consistent r.e. extension of PA,
$\sigma(x)$: numeration of T in T. Then

1. (Gödel, Feferman) If σ is Σ_1, then $T \nvdash \mathrm{Con}_\sigma$.

$\sigma(x), \sigma'(x)$: numerations of S and S' in T respectively.

Define

- $(\sigma|n)(x) :\equiv \sigma(x) \wedge x \leq \bar{n}$.
- $(\sigma \vee \sigma')(x) :\equiv \sigma(x) \vee \sigma'(x)$.

$(\sigma|n)(x)$ is a numeration of $\{\varphi \in S \mid \ulcorner \varphi \urcorner \leq n\}$.

$\mathrm{Con}_\sigma :\equiv \neg \mathrm{Pr}_\sigma(\ulcorner 0 = \bar{1} \urcorner)$.

Theorem

T: consistent r.e. extension of PA,
$\sigma(x)$: numeration of T in T. Then

1. (Gödel, Feferman) If σ is Σ_1, then $T \nvdash \mathrm{Con}_\sigma$.
2. (Mostowski) $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

Next we arithmetize the above construction of Henkin extension $T + Z$.

Next we arithmetize the above construction of Henkin extension $T + Z$.

- $C := \{c_n \mid n \in \omega\}$: set of new constants.
  $\mathcal{L}_C := \mathcal{L} \cup C$.

Next we arithmetize the above construction of Henkin extension $T + Z$.

- $C := \{c_n \mid n \in \omega\}$: **set of new constants.**
  $\mathcal{L}_C := \mathcal{L} \cup C$.
- **Extend Gödel numbering of $\mathcal{L}$-formulas to $\mathcal{L}_C$-formulas.**

Next we arithmetize the above construction of Henkin extension $T + Z$.

- $C := \{c_n \mid n \in \omega\}$: **set of new constants.**
  $\mathcal{L}_C := \mathcal{L} \cup C$.
- **Extend Gödel numbering of $\mathcal{L}$-formulas to $\mathcal{L}_C$-formulas.**
- $\mathrm{Fml}_C(x) \cdots$ "x **is an** $\mathcal{L}_C$**-formula**".
  $C(x) \cdots$ "x **is a new constant**".

Next we arithmetize the above construction of Henkin extension $T + Z$.

- $C := \{c_n \mid n \in \omega\}$: set of new constants.
  $\mathcal{L}_C := \mathcal{L} \cup C$.
- Extend Gödel numbering of $\mathcal{L}$-formulas to $\mathcal{L}_C$-formulas.
- $\mathrm{Fml}_C(x) \cdots$ "x is an $\mathcal{L}_C$-formula".
  $C(x) \cdots$ "x is a new constant".
- Define the p.r. set Z as above. Let $\zeta(x)$ be a suitable numeration of Z s.t.
  $\forall \varphi$, PA $\vdash \mathrm{Pr}_{\sigma \vee \zeta}(\ulcorner \varphi \urcorner) \to \mathrm{Pr}_{\sigma}(\ulcorner \varphi \urcorner)$.

Next we arithmetize the above construction of Henkin extension $T + Z$.

- $C := \{c_n \mid n \in \omega\}$: set of new constants.
  $\mathcal{L}_C := \mathcal{L} \cup C$.
- Extend Gödel numbering of $\mathcal{L}$-formulas to $\mathcal{L}_C$-formulas.
- $\mathrm{Fml}_C(x) \cdots$ "x is an $\mathcal{L}_C$-formula".
  $C(x) \cdots$ "x is a new constant".
- Define the p.r. set Z as above. Let $\zeta(x)$ be a suitable numeration of Z s.t.
  $\forall \varphi$, PA $\vdash \mathrm{Pr}_{\sigma \vee \zeta}(\ulcorner \varphi \urcorner) \to \mathrm{Pr}_{\sigma}(\ulcorner \varphi \urcorner)$.
- PA $\vdash \mathrm{Con}_\sigma \to \mathrm{Con}_{\sigma \vee \zeta}$.

Lastly, we arithmetize the Henkin completeness.

Lastly, we arithmetize the Henkin completeness.

For any $\mathcal{L}_A$-formula $\xi(x)$, define Hcm_ξ to be the conjunction of the following $\mathcal{L}_A$-sentences:

- $\forall x(\mathrm{Fml}_C(x) \to (\xi(\neg x) \leftrightarrow \neg\xi(x)))$;
- $\forall x, y(\mathrm{Fml}_C(x) \wedge \mathrm{Fml}_C(y) \to (\xi(x \vee y) \leftrightarrow (\xi(x) \vee \xi(y))))$;
- $\cdots$;
- $\forall x, y(\mathrm{Fml}_C(x) \to (\xi(\exists u x) \leftrightarrow \exists v(C(v) \wedge \xi(x[v/u]))))$.

Lastly, we arithmetize the Henkin completeness.

For any $\mathcal{L}_A$-formula $\xi(x)$, define Hcm_ξ to be the conjunction of the following $\mathcal{L}_A$-sentences:

- $\forall x(\mathrm{Fml}_C(x) \to (\xi(\neg x) \leftrightarrow \neg\xi(x)))$;
- $\forall x, y(\mathrm{Fml}_C(x) \land \mathrm{Fml}_C(y) \to (\xi(x \lor y) \leftrightarrow (\xi(x) \lor \xi(y))))$;
- $\cdots$;
- $\forall x, y(\mathrm{Fml}_C(x) \to (\xi(\exists u x) \leftrightarrow \exists v(C(v) \land \xi(x[v/u])))).$

Hcm_ξ states that the set defined by $\xi(x)$ is Henkin complete.

Lastly, we arithmetize the Henkin completeness.

For any $\mathcal{L}_A$-formula $\xi(x)$, define Hcm_ξ to be the conjunction of the following $\mathcal{L}_A$-sentences:

- $\forall x(\mathrm{Fml}_C(x) \to (\xi(\neg x) \leftrightarrow \neg \xi(x)))$;
- $\forall x, y(\mathrm{Fml}_C(x) \land \mathrm{Fml}_C(y) \to (\xi(x \lor y) \leftrightarrow (\xi(x) \lor \xi(y))))$;
- $\cdots$;
- $\forall x, y(\mathrm{Fml}_C(x) \to (\xi(\exists u x) \leftrightarrow \exists v(C(v) \land \xi(x[v/u]))))$.

Hcm_ξ states that the set defined by $\xi(x)$ is Henkin complete.

The arithmetized completeness theorem

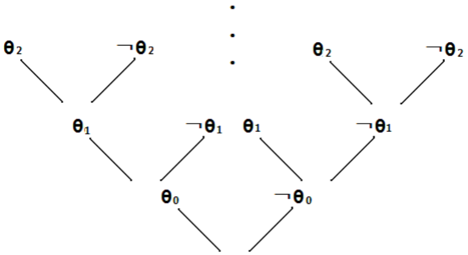
$\forall \sigma(x)$: **numeration of T, $\exists \xi(x)$: $\mathcal{L}_A$-formula s.t.**

1. **PA** $\vdash$ **Con**$_\sigma \to$ **Hcm**$_\xi$ **and**
2. **PA** $\vdash \forall x(\mathrm{Pr}_\sigma(x) \to \xi(x))$.

Proof.

- $\{\theta_n\}_{n\in\omega}$: p.r. enumeration of all $\mathcal{L}_C$-sentences.
- $\xi(x)\cdots$ "x is contained in the leftmost consistent path".
- PA $\vdash \forall x(\mathrm{Pr}_\sigma(x) \to \xi(x))$.
- PA $\vdash$ Con$_{\sigma\vee\zeta} \to$ Hcm$_\xi$.

□



Theorem(Feferman, 1960)

T: **extension of PA.** σ: **numeration of** S **in** T.

Then $S \leq T + \mathbf{Con}_{\sigma}$.

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathrm{Con}_\sigma$.

Proof.

- **$\xi(x)$: as in the arithmetized completeness theorem.**

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathrm{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \wedge \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathrm{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \wedge \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$
- By induction, $\forall \varphi$, $\mathrm{PA} + \mathrm{Hcm}_\xi \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathsf{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \wedge \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$
- By induction, $\forall \varphi$, $\mathsf{PA} + \mathsf{Hcm}_\xi \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- $\forall \varphi$, $T + \mathsf{Con}_\sigma \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathrm{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \wedge \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$
- By induction, $\forall \varphi$, $\mathrm{PA} + \mathrm{Hcm}_\xi \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- $\forall \varphi$, $T + \mathrm{Con}_\sigma \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- t is a translation of $\mathcal{L}_S$ into $\mathcal{L}_{T+\mathrm{Con}_\sigma}$.

Theorem(Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathsf{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \land \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$
- By induction, $\forall \varphi$, $\mathsf{PA} + \mathsf{Hcm}_\xi \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- $\forall \varphi$, $T + \mathsf{Con}_\sigma \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- t is a translation of $\mathcal{L}_S$ into $\mathcal{L}_{T+\mathsf{Con}_\sigma}$.
- φ: $\mathcal{L}_S$-sentence s.t. $S \vdash \varphi$. $T \vdash \mathrm{Pr}_\sigma(\ulcorner \varphi \urcorner)$.

Theorem (Feferman, 1960)

T: extension of PA. σ: numeration of S in T.

Then $S \leq T + \mathsf{Con}_\sigma$.

Proof.

- $\xi(x)$: as in the arithmetized completeness theorem.
- $d(x) :\equiv x = x$, $\eta_c(x) :\equiv C(x) \wedge \xi(\ulcorner c = \dot{x} \urcorner)$, $\cdots$
- By induction, $\forall \varphi$, $\mathsf{PA} + \mathsf{Hcm}_\xi \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- $\forall \varphi$, $T + \mathsf{Con}_\sigma \vdash t(\varphi) \leftrightarrow \xi(\ulcorner \varphi \urcorner)$.
- t is a translation of $\mathcal{L}_S$ into $\mathcal{L}_{T+\mathsf{Con}_\sigma}$.
- φ: $\mathcal{L}_S$-sentence s.t. $S \vdash \varphi$. $T \vdash \mathrm{Pr}_\sigma(\ulcorner \varphi \urcorner)$.
- $T + \mathsf{Con}_\sigma \vdash \xi(\ulcorner \varphi \urcorner)$. $T + \mathsf{Con}_\sigma \vdash t(\varphi)$.

□

We can construct a model by interpretation.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$**-structure** $\mathcal{N}$ **as follows:**

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}}$:= **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \wedge \eta_c(a)$;
- $\cdots$.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$**-structure** $\mathcal{N}$ **as follows:**

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}} :=$ **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \land \eta_c(a)$;
- $\cdots$.

By induction, $\forall \varphi$: $\mathcal{L}_S$**-sentence,** $\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{N} \models \varphi$.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$**-structure** $\mathcal{N}$ **as follows:**

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}}$:= **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \wedge \eta_c(a)$;
- $\cdots$.

By induction, $\forall \varphi$: $\mathcal{L}_S$**-sentence,** $\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{N} \models \varphi$.

- **Suppose** $S \vdash \varphi$.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$**-structure** $\mathcal{N}$ **as follows:**

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}}$:= **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \wedge \eta_c(a)$;
- $\cdots$.

By induction, $\forall \varphi$: $\mathcal{L}_S$**-sentence,** $\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{N} \models \varphi$.

- **Suppose** $S \vdash \varphi$.
- **Since** $S \leq T$, $T \vdash t(\varphi)$.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$-structure $\mathcal{N}$ as follows:

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}}$:= **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \wedge \eta_c(a)$;
- $\cdots$.

By induction, $\forall \varphi$: $\mathcal{L}_S$**-sentence,** $\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{N} \models \varphi$.

- **Suppose** $S \vdash \varphi$.
- **Since** $S \leq T$, $T \vdash t(\varphi)$.
- $\mathcal{M} \models t(\varphi)$, **so** $\mathcal{N} \models \varphi$.

We can construct a model by interpretation.

t: **interpretation of** S **in** T, $\mathcal{M}$: **model of** T.

Define an $\mathcal{L}_S$**-structure** $\mathcal{N}$ **as follows:**

- $|\mathcal{N}| := \{a \in |\mathcal{M}| \ : \ \mathcal{M} \models d(a)\}$;
- **For** $c \in \mathcal{L}_S$: **constant,**
  $c^{\mathcal{N}}$:= **the unique** $a \in |\mathcal{M}|$ **s.t.** $\mathcal{M} \models d(a) \wedge \eta_c(a)$;
- $\cdots$.

By induction, $\forall \varphi$: $\mathcal{L}_S$**-sentence,** $\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{N} \models \varphi$.

- **Suppose** $S \vdash \varphi$.
- **Since** $S \leq T$, $T \vdash t(\varphi)$.
- $\mathcal{M} \models t(\varphi)$, **so** $\mathcal{N} \models \varphi$.
- $\mathcal{N}$ **is a model of** S.

Definition

$\mathcal{M}, \mathcal{N}$: **models of arithmetic.**

$\mathcal{M}$ **is an initial segment of** $\mathcal{N}$ $(\mathcal{M} \subseteq_e \mathcal{N}) \overset{\text{def.}}{\Leftrightarrow}$

1. $|\mathcal{M}| \subseteq |\mathcal{N}|$ **and**
2. $\forall a \in |\mathcal{M}| \forall b \in |\mathcal{N}|, (\mathcal{N} \models b < a \Rightarrow b \in |\mathcal{M}|)$.

Definition

$\mathcal{M}, \mathcal{N}$: models of arithmetic.

$\mathcal{M}$ is an initial segment of $\mathcal{N}$ ($\mathcal{M} \subseteq_e \mathcal{N}$) $\overset{\text{def.}}{\Leftrightarrow}$

1. **$|\mathcal{M}| \subseteq |\mathcal{N}|$ and**
2. **$\forall a \in |\mathcal{M}| \forall b \in |\mathcal{N}|, (\mathcal{N} \models b < a \Rightarrow b \in |\mathcal{M}|)$.**

Theorem(Orey(1961), Hájek (1971,1972))

For any consistent r.e. extensions S, T of PA, T.F.A.E.:

(i) **$S \leq T$.**

(ii) **$\forall \mathcal{M} \models T \, \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.**

(iii) **$\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.**

(iv) **$\forall \sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.**

(i) $S \leq T$.

(ii) $\forall \mathcal{M} \models T \; \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(i) $\Rightarrow$ (ii)

- Let θ be an interpretation of S in T.
- Let $\mathcal{M}$ be any model of T.

(i) $S \leq T$.

(ii) $\forall \mathcal{M} \models T \ \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(i) $\Rightarrow$ (ii)

- Let θ be an interpretation of S in T.
- Let $\mathcal{M}$ be any model of T.
- Let $\mathcal{N}$ be a model of S defined by t and $\mathcal{M}$.

(i) $S \leq T$.

(ii) $\forall \mathcal{M} \models T \; \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(i) ⇒ (ii)

- Let θ be an interpretation of S in T.
- Let $\mathcal{M}$ be any model of T.
- Let $\mathcal{N}$ be a model of S defined by t and $\mathcal{M}$.
- Define a function f in $\mathcal{M}$ satisfying $f(0^{\mathcal{M}}) = 0^{\mathcal{N}}$ and $f(S^{\mathcal{M}}(a)) = S^{\mathcal{N}}(f(a))$.

(i) $S \leq T$.

(ii) $\forall \mathcal{M} \models T \ \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(i) $\Rightarrow$ (ii)

- Let θ be an interpretation of S in T.
- Let $\mathcal{M}$ be any model of T.
- Let $\mathcal{N}$ be a model of S defined by t and $\mathcal{M}$.
- Define a function f in $\mathcal{M}$ satisfying $f(0^{\mathcal{M}}) = 0^{\mathcal{N}}$ and $f(S^{\mathcal{M}}(a)) = S^{\mathcal{N}}(f(a))$.
- Then f is an isomorphism of an initial segment of $\mathcal{N}$.

(ii) $\forall \mathcal{M} \models T \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(ii) $\Rightarrow$ (iii)

- **Let θ be any Π_1 sentence s.t. $S \vdash \theta$.**
- **Let $\mathcal{M}$ be any model of T.**

(ii) $\forall \mathcal{M} \models T \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(ii) ⇒ (iii)

- **Let θ be any Π_1 sentence s.t. $S \vdash \theta$.**
- **Let $\mathcal{M}$ be any model of T.**
- **By (ii), $\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \models \mathcal{N}$.**

(ii) $\forall \mathcal{M} \models T \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(ii) $\Rightarrow$ (iii)

- Let θ be any Π_1 sentence s.t. $S \vdash \theta$.
- Let $\mathcal{M}$ be any model of T.
- By (ii), $\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \models \mathcal{N}$.
- $\mathcal{N} \models \theta$.

(ii) $\forall \mathcal{M} \models T \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(ii) $\Rightarrow$ (iii)

- Let θ be any Π_1 sentence s.t. $S \vdash \theta$.
- Let $\mathcal{M}$ be any model of T.
- By (ii), $\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \models \mathcal{N}$.
- $\mathcal{N} \models \theta$.
- Since θ is Π_1, $\mathcal{M} \models \theta$.

(ii) $\forall \mathcal{M} \models T \exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(ii) $\Rightarrow$ (iii)

- Let θ be any Π_1 sentence s.t. $S \vdash \theta$.
- Let $\mathcal{M}$ be any model of T.
- By (ii), $\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \models \mathcal{N}$.
- $\mathcal{N} \models \theta$.
- Since θ is Π_1, $\mathcal{M} \models \theta$.
- By completeness theorem, $T \vdash \theta$.

(i) $S \leq T$.

(iii) $\forall\theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(iv) $\forall\sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(iii) $\Rightarrow$ (iv)

- **By Mostowski's theorem,** $\forall n \in \omega$, $S \vdash \mathrm{Con}_{\sigma|n}$.

(i) $S \leq T$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(iv) $\forall \sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(iii) $\Rightarrow$ (iv)

- By Mostowski's theorem, $\forall n \in \omega$, $S \vdash \mathrm{Con}_{\sigma|n}$.
- By (iii), $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(i) $S \leq T$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(iv) $\forall \sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(iii) $\Rightarrow$ (iv)

- By Mostowski's theorem, $\forall n \in \omega$, $S \vdash \mathrm{Con}_{\sigma|n}$.
- By (iii), $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(iv) $\Rightarrow$ (i).

- Let $\sigma^*(x) :\equiv \sigma(x) \wedge \mathrm{Con}_{\sigma|x}$.

(i) $S \leq T$.

(iii) $\forall\theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(iv) $\forall\sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathsf{Con}_{\sigma|n}$.

(iii) $\Rightarrow$ (iv)

- By Mostowski's theorem, $\forall n \in \omega$, $S \vdash \mathsf{Con}_{\sigma|n}$.
- By (iii), $\forall n \in \omega$, $T \vdash \mathsf{Con}_{\sigma|n}$.

(iv) $\Rightarrow$ (i).

- Let $\sigma^*(x) :\equiv \sigma(x) \wedge \mathsf{Con}_{\sigma|x}$.
- Then $\sigma^*(x)$ numerates S in T and PA $\vdash \mathsf{Con}_{\sigma^*}$.

(i) $S \leq T$.

(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.

(iv) $\forall \sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$, $T \vdash \mathsf{Con}_{\sigma|n}$.

(iii) ⇒ (iv)

- By Mostowski's theorem, $\forall n \in \omega$, $S \vdash \mathsf{Con}_{\sigma|n}$.
- By (iii), $\forall n \in \omega$, $T \vdash \mathsf{Con}_{\sigma|n}$.

(iv) ⇒ (i).

- Let $\sigma^*(x) :\equiv \sigma(x) \wedge \mathsf{Con}_{\sigma|x}$.
- Then $\sigma^*(x)$ numerates S in T and PA $\vdash \mathsf{Con}_{\sigma^*}$.
- By Feferman's theorem, $S \leq T + \mathsf{Con}_{\sigma^*}$.

(i) $S \leq T$.
(iii) $\forall \theta$: Π_1 sentence, $S \vdash \theta \Rightarrow T \vdash \theta$.
(iv) $\forall \sigma(x)$: Σ_1 numeration of S $\forall n \in \omega$,
$T \vdash \mathrm{Con}_{\sigma|n}$.

(iii) $\Rightarrow$ (iv)

- By Mostowski's theorem, $\forall n \in \omega$, $S \vdash \mathrm{Con}_{\sigma|n}$.
- By (iii), $\forall n \in \omega$, $T \vdash \mathrm{Con}_{\sigma|n}$.

(iv) $\Rightarrow$ (i).

- Let $\sigma^*(x) :\equiv \sigma(x) \wedge \mathrm{Con}_{\sigma|x}$.
- Then $\sigma^*(x)$ numerates S in T and PA $\vdash \mathrm{Con}_{\sigma^*}$.
- By Feferman's theorem, $S \leq T + \mathrm{Con}_{\sigma^*}$.
- $S \leq T$.

□

- **Model theoretic proof of the second incompleteness theorem**
  (Kreisel)(Kikuchi,1994)

- **Model theoretic proof of the second incompleteness theorem (Kreisel)(Kikuchi,1994)**

Theorem

T, S: consistent r.e. extensions of PA. $\mathcal{M}$: models of T.
If $\mathcal{M} \models \mathrm{Con}_S$, then
$\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$ and
$\exists \xi(x)$: $\mathcal{L}_A$-formula s.t.

- $\forall \varphi$, $(\mathcal{M} \models \mathrm{Pr}_S(\ulcorner \varphi \urcorner) \Rightarrow \mathcal{N} \models \varphi)$,
- $\forall \varphi$, $(\mathcal{M} \models \xi(\ulcorner \varphi \urcorner) \Leftrightarrow \mathcal{N} \models \varphi)$.

- **Model theoretic proof of the second incompleteness theorem (Kreisel)(Kikuchi,1994)**

Theorem

T, S: consistent r.e. extensions of PA. $\mathcal{M}$: models of T.
If $\mathcal{M} \models \mathrm{Con}_S$, then
$\exists \mathcal{N} \models S$ s.t. $\mathcal{M} \subseteq_e \mathcal{N}$ and
$\exists \xi(x)$: $\mathcal{L}_A$-formula s.t.

- $\forall \varphi$, $(\mathcal{M} \models \mathrm{Pr}_S(\ulcorner \varphi \urcorner) \Rightarrow \mathcal{N} \models \varphi)$,
- $\forall \varphi$, $(\mathcal{M} \models \xi(\ulcorner \varphi \urcorner) \Leftrightarrow \mathcal{N} \models \varphi)$.

- **Suppose $\forall \mathcal{M} \models T$, $\mathcal{M} \models \mathrm{Con}_T$.**
- **By using Theorem, lead a contradiction.**

- **Faithful interpretability (Feferman, Kreisel, Orey, 1960)(Lindström, 1984)**

- **Faithful interpretability (Feferman, Kreisel, Orey, 1960)(Lindström, 1984)**

Definition

A interpretation t of S in T is faithful $\overset{\text{def.}}{\Leftrightarrow} \forall\varphi$, $(T \vdash t(\varphi) \Rightarrow S \vdash \varphi)$.

- **Faithful interpretability (Feferman, Kreisel, Orey, 1960)(Lindström, 1984)**

Definition

A interpretation t of S in T is faithful
$\overset{\text{def.}}{\Leftrightarrow}$ **$\forall \varphi$, $(T \vdash t(\varphi) \Rightarrow S \vdash \varphi)$.**
S is faithful interpretable in T
$\overset{\text{def.}}{\Leftrightarrow}$ **$\exists t$: faithful interpretation of S in T.**

- **Faithful interpretability**
  (Feferman, Kreisel, Orey, 1960)(Lindström, 1984)

Definition

A interpretation t of S in T is **faithful**
$\overset{\text{def.}}{\Leftrightarrow} \forall\varphi, (T \vdash t(\varphi) \Rightarrow S \vdash \varphi)$.
S is **faithful interpretable** in T
$\overset{\text{def.}}{\Leftrightarrow} \exists t$: faithful interpretation of S in T.

Theorem(Lindström)

S, T: r.e. extensions of PA. T.F.A.E.:

1. S is faithful interpretable in T.
2. $S \leq T$ and $\forall\varphi, (T \vdash \mathrm{Pr}_{\phi}(\ulcorner\varphi\urcorner) \Rightarrow S \vdash \varphi)$.
3. $\forall\theta : \Pi_1$ sentence, $(S \vdash \theta \Rightarrow T \vdash \theta)$ and
   $\forall\sigma : \Sigma_1$ sentence, $(T \vdash \sigma \Rightarrow S \vdash \theta)$ and

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: **extensions of PA.**

$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: **extensions of PA.**
$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \;\&\; T \leq S$.

- $\equiv$ **is an equivalence relation on extensions of PA.**

- Degrees of interpretability of extensions of PA (Lindström, 1979)

Definition

S, T: extensions of PA.

$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ is an equivalence relation on extensions of PA.
- Equivalence classes are called *degrees* of interpretability.

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: extensions of PA.
$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ **is an equivalence relation on extensions of PA.**
- **Equivalence classes are called** *degrees* **of interpretability.**
- $D_T :=$ **the set of degrees of extensions of** T**.**

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: **extensions of PA.**

$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ **is an equivalence relation on extensions of PA.**
- **Equivalence classes are called** *degrees* **of interpretability.**
- $D_T :=$ **the set of degrees of extensions of** T.
- $d(S) :=$ **degree of** S.

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: **extensions of PA.**
$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ **is an equivalence relation on extensions of PA.**
- **Equivalence classes are called** *degrees* **of interpretability.**
- $D_T :=$ **the set of degrees of extensions of** T.
- $d(S) :=$ **degree of** S.
- $d(S) \leq d(T) \overset{\text{def.}}{\Leftrightarrow} S \leq T$.

- **Degrees of interpretability of extensions of PA (Lindström, 1979)**

Definition

S, T: extensions of PA.
$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ is an equivalence relation on extensions of PA.
- Equivalence classes are called *degrees* of interpretability.
- $D_T :=$ the set of degrees of extensions of T.
- $d(S) :=$ degree of S.
- $d(S) \leq d(T) \overset{\text{def.}}{\Leftrightarrow} S \leq T$.
- $\mathcal{D}_T := (D_T, \leq)$.

- Degrees of interpretability of extensions of PA (Lindström, 1979)

Definition

S, T: extensions of PA.
$S \equiv T \overset{\text{def.}}{\Leftrightarrow} S \leq T \ \& \ T \leq S$.

- $\equiv$ is an equivalence relation on extensions of PA.
- Equivalence classes are called *degrees* of interpretability.
- $D_T :=$ the set of degrees of extensions of T.
- $d(S) :=$ degree of S.
- $d(S) \leq d(T) \overset{\text{def.}}{\Leftrightarrow} S \leq T$.
- $\mathcal{D}_T := (D_T, \leq)$.

Theorem(Lindström)

T is an r.e. consistent extension of PA, then $\mathcal{D}_T$ is a distributive lattice.

Open problem

S, T is Σ_1-sound r.e. extensions of PA. $\mathcal{D}_S \simeq \mathcal{D}_T$?

- **Interpretability logic (Visser, 1990)(Shavrukov, 1988)(Berarducci, 1990)**

- **Interpretability logic (Visser, 1990)(Shavrukov, 1988)(Berarducci, 1990)**

$\mathrm{Int}_T(x, y) \cdots$ " $T + x$ **is interpretable in** $T + y$**.**

- **Interpretability logic (Visser, 1990)(Shavrukov, 1988)(Berarducci, 1990)**

$\mathrm{Int}_T(x, y) \cdots$ " $T + x$ **is interpretable in** $T + y$.

Propositional modal logic **ILM**

ILM $=$ **GL**$+$ **the following axioms:**

- $\Box(A \to B) \to A \rhd B$;
- $A \rhd B \land B \rhd C \to A \rhd C$;
- $A \rhd C \land B \rhd C \to (A \lor B) \rhd C$;
- $A \rhd B \to (\Diamond A \to \Diamond B)$;
- $\Diamond A \rhd A$;
- $A \rhd B \to ((A \land \Box C) \rhd (B \land \Box C))$.

- **Interpretability logic (Visser, 1990)(Shavrukov, 1988)(Berarducci, 1990)**

$\mathrm{Int}_T(x, y) \cdots$ " $T + x$ **is interpretable in** $T + y$.

Propositional modal logic ILM

ILM $=$ **GL**$+$ **the following axioms:**

- $\Box(A \to B) \to A \rhd B$;
- $A \rhd B \land B \rhd C \to A \rhd C$;
- $A \rhd C \land B \rhd C \to (A \lor B) \rhd C$;
- $A \rhd B \to (\Diamond A \to \Diamond B)$;
- $\Diamond A \rhd A$;
- $A \rhd B \to ((A \land \Box C) \rhd (B \land \Box C))$.

Theorem(Shavrukov,Berarducci)

$\forall A$, (**ILM** $\vdash A \Leftrightarrow A$ **is arithmetically valid**).

References

- **S. Feferman.** *Transfinite recursive progressions of axiomatic theories*. **J. Symbolic Logic 27 (1962) 259–316.**
- **R. Kaye.** *Models of Peano arithmetic*, **Oxford Logic Guides, 15. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1991.**
- **P. Lindström.** *Aspects of incompleteness*. **Lecture Notes in Logic, 10. Springer-Verlag, Berlin, 1997.**
- **S. Orey.** *Relative interpretations*. **Z. Math. Logik Grundlagen Math. 7 (1961) 146–153.**
- **C. Smoryński.** *The incompleteness theorems*. **Handbook of Mathematical Logic (J. Barwise, ed), North-Holland, Amsterdam (1977), 821–865.**
- **A. Tarski.** *Undecidable theories*. **In collaboration with Andrzej Mostowski and Raphael M. Robinson. Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Company, Amsterdam, 1953.**