Full Satisfaction Classes in a General Setting

Albert Visser + Ali Enayat

Model Theory and Proof Theory of Arithemtic
A Memorial Conference in Honor of Henryk Kotlarski and Zygmunt Ratajczyk

July 24, 2012, Bedlewo

On Tarski(S, F)

On Tarski(S, F)

- Given a base theory B we wish to define certain canonical associated *satisfaction* theories here denoted B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$, all of which are formulated in an *expansion* of the language $\mathcal{L}_{\mathsf{B}}$ by adding a new *binary* predicate $\mathsf{S}(x, y)$.

On Tarski(S, F)

- Given a base theory B we wish to define certain canonical associated *satisfaction* theories here denoted B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$, all of which are formulated in an *expansion* of the language $\mathcal{L}_{\mathsf{B}}$ by adding a new *binary* predicate $\mathsf{S}(x, y)$.
- Tarski(S, F) consists of the following axioms:

On Tarski(S, F)

- Given a base theory B we wish to define certain canonical associated *satisfaction* theories here denoted B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$, all of which are formulated in an *expansion* of the language $\mathcal{L}_{\mathsf{B}}$ by adding a new *binary* predicate $\mathsf{S}(x, y)$.
- Tarski(S, F) consists of the following axioms:
- $\mathsf{tarski}_0(\mathsf{S}, \mathsf{F}) :=$
  $(\mathsf{F}(x) \rightarrow \mathsf{Form}(x)) \wedge (\mathsf{S}(x, \alpha) \rightarrow \mathsf{F}(x) \wedge \mathsf{Asn}(\alpha, x))$.

On Tarski(S, F)

- Given a base theory B we wish to define certain canonical associated *satisfaction* theories here denoted B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$, all of which are formulated in an *expansion* of the language $\mathcal{L}_{\mathsf{B}}$ by adding a new *binary* predicate $\mathsf{S}(x, y)$.
- Tarski(S, F) consists of the following axioms:
- $\mathsf{tarski}_0(\mathsf{S}, \mathsf{F}) :=$
  $(\mathsf{F}(x) \to \mathsf{Form}(x)) \wedge (\mathsf{S}(x, \alpha) \to \mathsf{F}(x) \wedge \mathsf{Asn}(\alpha, x))$.
- $\mathsf{tarski}_{1,R}(\mathsf{S}, \mathsf{F}) :=$

$$(\mathsf{F}(x) \wedge (x = \ulcorner R\,(t_0, \cdots, t_{n-1})\urcorner) \wedge \mathsf{Asn}(\alpha, x)) \to$$
$$(\mathsf{S}(x, \alpha) \leftrightarrow R\left([\alpha]_{t_0} \cdots, [\alpha]_{t_{n-1}}\right)).$$

On Tarski(S, F), continued

On Tarski(S, F), continued

- $\mathsf{tarski}_2(\mathsf{S},\mathsf{F}) := \left(\begin{array}{c} \mathsf{F}(x) \wedge (x = \ulcorner \neg y \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{array}\right) \rightarrow (\mathsf{S}(x,\alpha) \leftrightarrow \neg \mathsf{S}(y,\alpha))$.

On Tarski(S, F), continued

- $\mathsf{tarski}_2(\mathsf{S},\mathsf{F}) := \begin{pmatrix} \mathsf{F}(x) \wedge (x = \ulcorner \neg y \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{pmatrix} \rightarrow (\mathsf{S}(x,\alpha) \leftrightarrow \neg \mathsf{S}(y,\alpha))\,.$

- $\mathsf{tarski}_3(\mathsf{S},\mathsf{F}) :=$

$$\begin{pmatrix} \mathsf{F}(x) \wedge (x = \ulcorner y_1 \vee y_2 \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{pmatrix} \rightarrow \left(\mathsf{S}(x,\alpha) \leftrightarrow \begin{pmatrix} \mathsf{S}\,(y_1, \alpha \restriction \mathsf{FV}(y_1)) \vee \\ \mathsf{S}\,(y_2, \alpha \restriction \mathsf{FV}(y_2)) \end{pmatrix}\right).$$

On Tarski(S, F), continued

- $\mathsf{tarski}_2(\mathsf{S},\mathsf{F}) := \begin{pmatrix} \mathsf{F}(x) \wedge (x = \ulcorner \neg y \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{pmatrix} \rightarrow (\mathsf{S}(x,\alpha) \leftrightarrow \neg \mathsf{S}(y,\alpha))$.

- $\mathsf{tarski}_3(\mathsf{S},\mathsf{F}) :=$

$$\begin{pmatrix} \mathsf{F}(x) \wedge (x = \ulcorner y_1 \vee y_2 \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{pmatrix} \rightarrow \left(\mathsf{S}(x,\alpha) \leftrightarrow \begin{pmatrix} \mathsf{S}\left(y_1, \alpha \restriction \mathsf{FV}(y_1)\right) \vee \\ \mathsf{S}\left(y_2, \alpha \restriction \mathsf{FV}(y_2)\right) \end{pmatrix}\right).$$

- $\mathsf{tarski}_4(\mathsf{S},\mathsf{F}) :=$

$$\begin{pmatrix} \mathsf{F}(x) \wedge (x = \ulcorner \exists v_i \; y \urcorner) \wedge \\ \mathsf{Asn}(\alpha, x) \end{pmatrix} \rightarrow \begin{pmatrix} \mathsf{S}(x,\alpha) \leftrightarrow \\ \exists \alpha' \left(\alpha \sim_{v_i} \alpha' \wedge \mathsf{S}(y,\alpha')\right) \end{pmatrix}.$$

B^{FS}, B^{IS}, and B^{FIS}

B^{FS}, B^{IS}, and B^{FIS}

- $B^{FS} := B \cup \text{Tarski}(S, \text{Form})$.

B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$

- $\mathsf{B}^{\mathsf{FS}} := \mathsf{B} \cup \mathsf{Tarski}(\mathsf{S}, \mathsf{Form})$.
- $\mathsf{B}^{\mathsf{IS}} := \mathsf{B} \cup \{\mathsf{Tarski}(\mathsf{S}, \mathsf{Form}_n) : n \in \omega\} \cup \mathsf{Ind}(\mathsf{S})$, where Form_n is the collection of formulas $\mathcal{L}_\mathsf{B}$ with quantifier alternation depth at most n and $\mathsf{Ind}(\mathsf{S})$ is the full scheme of induction on $\mathbb{N}$ in the language $\mathcal{L}_\mathsf{B}(\mathsf{S}) := \mathcal{L}_\mathsf{B} \cup \{\mathsf{S}\}$.

B^{FS}, B^{IS}, and $\mathsf{B}^{\mathsf{FIS}}$

- $\mathsf{B}^{\mathsf{FS}} := \mathsf{B} \cup \mathsf{Tarski}\,(\mathsf{S}, \mathsf{Form})$.
- $\mathsf{B}^{\mathsf{IS}} := \mathsf{B} \cup \{\mathsf{Tarski}\,(\mathsf{S}, \mathsf{Form}_n) : n \in \omega\} \cup \mathsf{Ind}(\mathsf{S})$, where Form_n is the collection of formulas $\mathcal{L}_{\mathsf{B}}$ with quantifier alternation depth at most n and $\mathsf{Ind}(\mathsf{S})$ is the full scheme of induction on $\mathbb{N}$ in the language $\mathcal{L}_{\mathsf{B}}(\mathsf{S}) := \mathcal{L}_{\mathsf{B}} \cup \{\mathsf{S}\}$.
- $\mathsf{B}^{\mathsf{FIS}} := \mathsf{B}^{\mathsf{FS}} \cup \mathsf{Ind}(\mathsf{S}) = \mathsf{B}^{\mathsf{FS}} \cup \mathsf{B}^{\mathsf{IS}}$.

Satisfaction Classes

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathrm{B}$, $F \subseteq \mathrm{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathsf{B}$, $F \subseteq \mathsf{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.
- S is an *F-satisfaction class* if $(\mathcal{M}, S, F) \models \mathsf{Tarski(S, F)}$.

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathsf{B}$, $F \subseteq \mathsf{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.
- S is an *F-satisfaction class* if $(\mathcal{M}, S, F) \models \mathsf{Tarski(S, F)}$.
- *If* $\mathsf{F{=}Form}^{\mathcal{M}} \cap \ \omega$, then we say that F is the set of *standard* $\mathcal{L}_{\mathsf{B}}$-formulas of $\mathcal{M}$. In this case there is a unique F-satisfaction class on $\mathcal{M}$, which we refer to as the *Tarskian satisfaction class on $\mathcal{M}$.*

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathsf{B}$, $F \subseteq \mathrm{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.
- S is an *F-satisfaction class* if $(\mathcal{M}, S, F) \models \mathrm{Tarski(S, F)}$.
- *If* F=$\mathrm{Form}^{\mathcal{M}} \cap\ \omega$, then we say that F is the set of *standard* $\mathcal{L}_{\mathsf{B}}$-formulas of $\mathcal{M}$. In this case there is a unique F-satisfaction class on $\mathcal{M}$, which we refer to as the *Tarskian satisfaction class on* $\mathcal{M}$.
- S is a *full satisfaction class on* $\mathcal{M}$ if S is an *F-satisfaction class* for $F := \mathrm{Form}^{\mathcal{M}}$. This is equivalent to $(\mathcal{M}, S) \models \mathsf{B}^{\mathsf{FS}}$.

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathrm{B}$, $F \subseteq \mathrm{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.
- S is an *F-satisfaction class* if $(\mathcal{M}, S, F) \models \mathrm{Tarski(S, F)}$.
- *If* $\mathrm{F{=}Form}^{\mathcal{M}} \cap \ \omega$, then we say that F is the set of *standard* $\mathcal{L}_{\mathrm{B}}$-formulas of $\mathcal{M}$. In this case there is a unique F-satisfaction class on $\mathcal{M}$, which we refer to as the *Tarskian satisfaction class on $\mathcal{M}$.*
- S is a *full satisfaction class on $\mathcal{M}$* if S is an *F-satisfaction class* for $F := \mathrm{Form}^{\mathcal{M}}$. This is equivalent to $(\mathcal{M}, S) \models \mathrm{B}^{\mathrm{FS}}$.
- S is a *satisfaction class on $\mathcal{M}$*, if either **(i)** $\mathcal{M}$ is ω-standard and S is the usual Tarskian satisfaction relation $Sat_{\mathcal{M}}$ on $\mathcal{M}$; **or** $\mathcal{M}$ is **not** ω-standard and there is some *nonstandard* integer c of $\mathcal{M}$ such that the expansion $(\mathcal{M}, S, F_{\leq c}) \models \mathrm{Tarski(S, F)}$.

Satisfaction Classes

- Suppose $\mathcal{M} \models \mathrm{B}$, $F \subseteq \mathrm{Form}^{\mathcal{M}}$, where F is *closed under direct subformulas*, and let S be a binary relation on M.
- S is an *F-satisfaction class* if $(\mathcal{M}, S, F) \models \mathrm{Tarski(S, F)}$.
- *If* $\mathrm{F{=}Form}^{\mathcal{M}} \cap \ \omega$, then we say that F is the set of *standard* $\mathcal{L}_{\mathrm{B}}$-formulas of $\mathcal{M}$. In this case there is a unique F-satisfaction class on $\mathcal{M}$, which we refer to as the *Tarskian satisfaction class on $\mathcal{M}$.*
- S is a *full satisfaction class on $\mathcal{M}$* if S is an *F-satisfaction class* for $F := \mathrm{Form}^{\mathcal{M}}$. This is equivalent to $(\mathcal{M}, S) \models \mathrm{B}^{\mathrm{FS}}$.
- S is a *satisfaction class on $\mathcal{M}$*, if either **(i)** $\mathcal{M}$ is ω-standard and S is the usual Tarskian satisfaction relation $Sat_{\mathcal{M}}$ on $\mathcal{M}$; **or** $\mathcal{M}$ is **not** ω-standard and there is some *nonstandard* integer c of $\mathcal{M}$ such that the expansion $(\mathcal{M}, S, F_{\leq c}) \models \mathrm{Tarski(S, F)}$.
- A satisfaction class S on $\mathcal{M}$ is said to be an *inductive satisfaction class* on $\mathcal{M}$ if $(\mathcal{M}, S) \models \mathrm{B}^{\mathrm{IS}}$.

Strongly Reflexive Base Theories (1)

Strongly Reflexive Base Theories (1)

- A base theory B is *strongly reflexive* if B is bi-interpretable with the theory T_{B} formulated in the language of set theory $\{\in\}$ such that T satisfies the following two properties; note that property (a) implies that B is an *inductive* base theory.

  (a) $T_{\mathrm{B}} \vdash \mathrm{KP} + \mathrm{Ind} + \mathrm{Infinity}$, where KP is Kripke-Platek set theory.

  (b) For each *sentence* φ in the language of set theory, T_{B} proves the implication

$$\varphi \to \exists x\ \varphi^{(x)},$$

  where x does not occur in φ and $\varphi^{(x)}$ is the formula obtained by relativizing all of the quantifiers of φ to x.

Strongly Reflexive Base Theories (2)

Strongly Reflexive Base Theories (2)

- **Examples.** *Any extension (in the same language) of the following theories is strongly reflexive.*

Strongly Reflexive Base Theories (2)

- **Examples.** *Any extension* (*in the same language*) *of the following theories is strongly reflexive.*
- *Zermelo-Fraenkel set theory* ZF.

Strongly Reflexive Base Theories (2)

- **Examples.** *Any extension (in the same language) of the following theories is strongly reflexive.*
- *Zermelo-Fraenkel set theory* ZF.
- *Second Order Arithmetic* Z_2 *augmented with the full scheme* Π^1_∞-DC *of dependent choice.*

- **Examples.** *Any extension (in the same language) of the following theories is strongly reflexive.*
- *Zermelo-Fraenkel set theory* ZF.
- *Second Order Arithmetic* Z_2 *augmented with the full scheme* Π^1_∞-DC *of dependent choice.*
- *Kelley-Morse theory of classes* KM *augmented with the full scheme* Π^1_∞-DC *of dependent choice.*

Strongly Reflexive Base Theories (2)

- **Examples.** *Any extension (in the same language) of the following theories is strongly reflexive.*
- *Zermelo-Fraenkel set theory* ZF.
- *Second Order Arithmetic* Z_2 *augmented with the full scheme* Π^1_∞-DC *of dependent choice.*
- *Kelley-Morse theory of classes* KM *augmented with the full scheme* Π^1_∞-DC *of dependent choice.*
- **Theorem.** *Every model* $\mathcal{M}$ *of a strongly reflexive theory base theory* B *is elementarily equivalent to a model* $\mathcal{N}$ *that carries a full* **inductive** *satisfaction class* S.

Strongly Reflexive Base Theories (3)

Strongly Reflexive Base Theories (3)

- **Corollary.** *Every countable recursively saturated model of a strongly reflexive base theory carries an inductive full satisfaction class.*

Strongly Reflexive Base Theories (3)

- **Corollary.** *Every countable recursively saturated model of a strongly reflexive base theory carries an inductive full satisfaction class.*
- **Corollary** (Conservativity Results).
  (a) $\mathsf{B}^{\mathsf{FS}} + \mathsf{Ind}(\mathsf{S})$ *is conservative over* B *for every strongly reflexive base theory* B.

  (b) $\mathsf{ZF}^{\mathsf{FS}} + \mathsf{Sep}(\mathsf{S})$ *is a conservative extension of* ZF.

The Core Construction (1)

The Core Construction (1)

- **Core Lemma.** *Let* $\mathcal{N}_0 \models \mathrm{B}$ *and suppose* S_0 *is an* F_0*-satisfaction class, where* $F_0 \subseteq F_1 := \mathrm{Form}^{\mathcal{N}_0}$*.* **Then** *there is an elementary extension* $\mathcal{N}_1$ *of* $\mathcal{N}_0$ *that carries an* F_1*-satisfaction class* $S_1 \supseteq S_0$.

The Core Construction (1)

- **Core Lemma.** *Let* $\mathcal{N}_0 \models \mathrm{B}$ *and suppose* S_0 *is an* F_0*-satisfaction class, where* $F_0 \subseteq F_1 := \mathrm{Form}^{\mathcal{N}_0}$*.* **Then** *there is an elementary extension* $\mathcal{N}_1$ *of* $\mathcal{N}_0$ *that carries an* F_1*-satisfaction class* $S_1 \supseteq S_0$.
- **Proof:** Let $\mathcal{L}^{+}_{\mathrm{B}}(\mathcal{N}_0)$ be the language obtained by enriching $\mathcal{L}_{\mathrm{B}}$ with constant symbols for each member of N_0, and new *unary* predicates U_c for each $c \in \mathrm{Form}^{\mathcal{N}_0}$.

The Core Construction (2)

The Core Construction (2)

- If $R \in \mathcal{L}_{\mathrm{B}}$ and $\mathcal{N}_0 \models c = \ulcorner R(t_0, \cdots, t_{n-1}) \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge R([\alpha]_{t_0}, \cdots, [\alpha]_{t_{n-1}}) \right).$$

The Core Construction (2)

- If $R \in \mathcal{L}_{\mathrm{B}}$ and $\mathcal{N}_0 \models c = \ulcorner R(t_0, \cdots, t_{n-1}) \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge R([\alpha]_{t_0}, \cdots, [\alpha]_{t_{n-1}}) \right).$$

- If $\mathcal{N} \models c = \ulcorner \neg d \urcorner$ then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge \neg \mathsf{U}_d(\alpha) \right).$$

The Core Construction (2)

- If $R \in \mathcal{L}_\mathsf{B}$ and $\mathcal{N}_0 \models c = \ulcorner R(t_0, \cdots, t_{n-1}) \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge R([\alpha]_{t_0}, \cdots, [\alpha]_{t_{n-1}}) \right).$$

- If $\mathcal{N} \models c = \ulcorner \neg d \urcorner$ then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge \neg \mathsf{U}_d(\alpha) \right).$$

- If $\mathcal{N} \models c = \ulcorner d_1 \vee d_2 \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge (\mathsf{U}_{d_1}(\alpha \restriction \mathsf{FV}(d_1)) \vee \mathsf{U}_{d_2}(\alpha \restriction \mathsf{FV}(d_2))) \right).$$

The Core Construction (2)

- If $R \in \mathcal{L}_{\mathrm{B}}$ and $\mathcal{N}_0 \models c = \ulcorner R(t_0, \cdots, t_{n-1}) \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge R([\alpha]_{t_0}, \cdots, [\alpha]_{t_{n-1}}) \right).$$

- If $\mathcal{N} \models c = \ulcorner \neg d \urcorner$ then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge \neg \mathsf{U}_d(\alpha) \right).$$

- If $\mathcal{N} \models c = \ulcorner d_1 \vee d_2 \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge \left(\mathsf{U}_{d_1}(\alpha \restriction \mathsf{FV}(d_1)) \vee \mathsf{U}_{d_2}(\alpha \restriction \mathsf{FV}(d_2)) \right) \right).$$

- If $\mathcal{N} \models c = \ulcorner \exists v_a \ b \urcorner$, then

$$\theta_c := \forall \alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \exists \alpha' \left(\alpha \sim_{v_a} \alpha' \wedge \ \mathsf{U}_b(\alpha') \right) \right).$$

The Core Construction (2)

- If $R \in \mathcal{L}_{\mathrm{B}}$ and $\mathcal{N}_0 \models c = \ulcorner R(t_0, \cdots, t_{n-1}) \urcorner$, then
$$\theta_c := \forall\alpha \left(\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge R([\alpha]_{t_0}, \cdots, [\alpha]_{t_{n-1}}) \right).$$
- If $\mathcal{N} \models c = \ulcorner \neg d \urcorner$ then
$$\theta_c := \forall\alpha \, (\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge \neg \mathsf{U}_d(\alpha)) \,.$$
- If $\mathcal{N} \models c = \ulcorner d_1 \vee d_2 \urcorner$, then
$$\theta_c := \forall\alpha \, (\mathsf{U}_c(\alpha) \leftrightarrow \alpha \in A_c \wedge (\mathsf{U}_{d_1}(\alpha \restriction \mathsf{FV}(d_1)) \vee \mathsf{U}_{d_2}(\alpha \restriction \mathsf{FV}(d_2)))) \,.$$
- If $\mathcal{N} \models c = \ulcorner \exists v_a \; b \urcorner$, then
$$\theta_c := \forall\alpha \, (\mathsf{U}_c(\alpha) \leftrightarrow \exists\alpha' \, (\alpha \sim_{v_a} \alpha' \wedge \; \mathsf{U}_b(\alpha'))) \,.$$
- Let $\Gamma := \{\mathsf{U}_c(\alpha) : c \in F_0 \text{ and } (c, \alpha) \in S_0\}$ and define
$$\mathsf{Th}^+(\mathcal{N}_0) := \mathsf{Th}(\mathcal{N}_0, c)_{c \in N_0} \cup \Theta \cup \Gamma.$$

The Core Construction (3)

- We now proceed to show that $\mathrm{Th}^+(\mathcal{N}_0)$ is consistent by demonstrating that each finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ is interpretable in $(\mathcal{N}_0, S_0)$.

The Core Construction (3)

- We now proceed to show that $\mathrm{Th}^+(\mathcal{N}_0)$ is consistent by demonstrating that each finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ is interpretable in $(\mathcal{N}_0, S_0)$.
- To this end, suppose T_0 is a finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ and let C consist of the collection of constants c that appear in at least one of the sentences in $T_0 \cap \Theta$. If $C = \emptyset$, T_0 is readily seen to be consistent, so we shall assume that $C \neq \emptyset$ for the rest of the argument.

The Core Construction (3)

- We now proceed to show that $\mathrm{Th}^+(\mathcal{N}_0)$ is consistent by demonstrating that each finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ is interpretable in $(\mathcal{N}_0, S_0)$.
- To this end, suppose T_0 is a finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ and let C consist of the collection of constants c that appear in at least one of the sentences in $T_0 \cap \Theta$. If $C = \emptyset$, T_0 is readily seen to be consistent, so we shall assume that $C \neq \emptyset$ for the rest of the argument.
- Our goal is to construct subsets $\{U_c : c \in C\}$ of N_0 such that the following two conditions hold when U_c, is interpreted by U_c:

  (1) $(\mathcal{N}_0, U_c)_{c\in C} \models \{\theta_c : c \in C\}$ and

  (2) $\alpha \in U_c$ whenever $c \in C \cap F_0$ and $(c, \alpha) \in S_0$.

The Core Construction (3)

- We now proceed to show that $\mathrm{Th}^+(\mathcal{N}_0)$ is consistent by demonstrating that each finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ is interpretable in $(\mathcal{N}_0, S_0)$.
- To this end, suppose T_0 is a finite subset of $\mathrm{Th}^+(\mathcal{N}_0)$ and let C consist of the collection of constants c that appear in at least one of the sentences in $T_0 \cap \Theta$. If $C = \emptyset$, T_0 is readily seen to be consistent, so we shall assume that $C \neq \emptyset$ for the rest of the argument.
- Our goal is to construct subsets $\{U_c : c \in C\}$ of N_0 such that the following two conditions hold when U_c, is interpreted by U_c:

  (1) $(\mathcal{N}_0, U_c)_{c \in C} \models \{\theta_c : c \in C\}$ and

  (2) $\alpha \in U_c$ whenever $c \in C \cap F_0$ and $(c, \alpha) \in S_0$.
- We shall construct $\{U_c : c \in C\}$ *in stages*, beginning with the simplest formulas in C, and working our way up using Tarski rules for more complex ones.

The Core Construction (4)

The Core Construction (4)

- Let $c \lhd d$ express "c is a direct subformula of d".

The Core Construction (4)

- Let $c \lhd d$ express "c is a direct subformula of d".
- Define $\lhd^*$ on C by:

$$c \lhd^* d \text{ iff } (c \lhd d)^{\mathcal{N}_0} \text{ and } \theta_d \in T_0 \cap \Theta.$$

The Core Construction (5)

The Core Construction (5)

- The finiteness of C implies that $(C, \lhd^*)$ is well-founded, which in turn helps us define a useful measure of complexity for $c \in C$ using the following recursive definition:

The Core Construction (5)

- The finiteness of C implies that $(C, \lhd^*)$ is well-founded, which in turn helps us define a useful measure of complexity for $c \in C$ using the following recursive definition:
  - $\mathsf{rank}_C(x) := \sup\{\mathsf{rank}_C(y) + 1 : x \in C \text{ and } (y \lhd^* x)^{\mathcal{N}_0}\}$.

The Core Construction (5)

- The finiteness of C implies that $(C, \lhd^*)$ is well-founded, which in turn helps us define a useful measure of complexity for $c \in C$ using the following recursive definition:
  - $\mathsf{rank}_C(x) := \sup\{\mathsf{rank}_C(y) + 1 : x \in C \text{ and } (y \lhd^* x)^{\mathcal{N}_0}\}$.
- Note that $\mathsf{rank}_G(c) = 0$ precisely when there is no $x \in C$ such that $(x \lhd^* c)^{\mathcal{N}_0}$.

The Core Construction (5)

- The finiteness of C implies that $(C, \lhd^*)$ is well-founded, which in turn helps us define a useful measure of complexity for $c \in C$ using the following recursive definition:
  - $\text{rank}_C(x) := \sup\{\text{rank}_C(y) + 1 : x \in C \text{ and } (y \lhd^* x)^{\mathcal{N}_0}\}$.
- Note that $\text{rank}_G(c) = 0$ precisely when there is no $x \in C$ such that $(x \lhd^* c)^{\mathcal{N}_0}$.
- Next,let

$$C_i := \{x \in C : \text{rank}_C(x)\} \leq i\}.$$

The Core Construction (5)

- The finiteness of C implies that $(C, \lhd^*)$ is well-founded, which in turn helps us define a useful measure of complexity for $c \in C$ using the following recursive definition:
  - $\text{rank}_C(x) := \sup\{\text{rank}_C(y) + 1 : x \in C \text{ and } (y \lhd^* x)^{\mathcal{N}_0}\}$.
- Note that $\text{rank}_G(c) = 0$ precisely when there is no $x \in C$ such that $(x \lhd^* c)^{\mathcal{N}_0}$.
- Next,let

$$C_i := \{x \in C : \text{rank}_C(x)\} \leq i\}.$$

- Observe that since C is finite, $C_0 \neq \emptyset$, and $c \in C_0$ iff $c \in C$ and C does not contain the code of any subformula of the formula coded by c. Moreover, if $c \in C_{i+1}$, then the codes of every immediate subformula of the formula coded by c are in C_i. This observation ensures that the following recursive clauses yield a well-defined U_c for each $c \in C$.

The Core Construction (6)

The Core Construction (6)

- If $c \in C_0$ then $U_c := \begin{cases} \{\alpha : (c, \alpha) \in S_0\}, \text{ if } c \in F_0; \\ U_c := \emptyset, \text{ if } c \notin F_0. \end{cases}$

The Core Construction (6)

- If $c \in C_0$ then $U_c := \begin{cases} \{\alpha : (c, \alpha) \in S_0\}, \text{ if } c \in F_0; \\ U_c := \emptyset, \text{ if } c \notin F_0. \end{cases}$
- If $c \in C_{i+1} \setminus C_i$ and $\ulcorner c = \neg d \urcorner$, then

$$U_c := \{\alpha \in A_c : \alpha \notin U_d\}.$$

The Core Construction (6)

- If $c \in C_0$ then $U_c := \begin{cases} \{\alpha : (c, \alpha) \in S_0\}, \text{ if } c \in F_0; \\ U_c := \emptyset, \text{ if } c \notin F_0. \end{cases}$
- If $c \in C_{i+1} \setminus C_i$ and $\ulcorner c = \neg d \urcorner$, then

$$U_c := \{\alpha \in A_c : \alpha \notin U_d\}.$$

- If $c \in C_{i+1} \setminus C_i$ and $c = \ulcorner a \vee b \urcorner$, then

$$U_c := \{\alpha \in A_c : \alpha \restriction \mathsf{FV}(a) \in U_a \text{ or } \alpha \restriction \mathsf{FV}(b) \in U_b\}.$$

The Core Construction (6)

- If $c \in C_0$ then $U_c := \begin{cases} \{\alpha : (c,\alpha) \in S_0\}, \text{ if } c \in F_0; \\ U_c := \emptyset, \text{ if } c \notin F_0. \end{cases}$

- If $c \in C_{i+1} \backslash C_i$ and $\ulcorner c = \neg d \urcorner$, then

$$U_c := \{\alpha \in A_c : \alpha \notin U_d\}.$$

- If $c \in C_{i+1} \backslash C_i$ and $c = \ulcorner a \vee b \urcorner$, then

$$U_c := \{\alpha \in A_c : \alpha \restriction \mathsf{FV}(a) \in U_a \text{ or } \alpha \restriction \mathsf{FV}(b) \in U_b\}.$$

- If $c \in C_{i+1} \backslash C_i$ and $c = \ulcorner \exists v_a \; b \urcorner$, then

$$U_c := \left\{\alpha \in A_c : \exists \alpha' \in N \; (\alpha \sim_{v_a} \alpha' \text{ and } \alpha \in U_b\right\}.$$

The Core Construction (7)

The Core Construction (7)

- **Core Theorem.** *Let $\mathcal{M}_0$ be a model of* B *of any cardinality. There is an elementary extension $\mathcal{M}$ of $\mathcal{M}_0$ that admits a full satisfaction class.*

The Core Construction (7)

- **Core Theorem.** *Let $\mathcal{M}_0$ be a model of* B *of any cardinality. There is an elementary extension $\mathcal{M}$ of $\mathcal{M}_0$ that admits a full satisfaction class.*

- **Proof:** Let F_0 be the set of *atomic* $\mathcal{N}$-formulas and let S_0 be the obvious satisfaction predicate for F_0. Then by the Lemma there is an elementary extension $\mathcal{M}_1$ of $\mathcal{M}_0$ that carries a $\mathrm{Form}_{\mathrm{B}}^{\mathcal{M}_0}$ satisfaction class. Thanks to Lemma 3.1, this argument can be carried out countably many times to yield two sequences $\langle \mathcal{M}_i : i \in \omega \rangle$ and $\langle S_i : i \in \omega \rangle$ that satisfy the following two properties:

The Core Construction (7)

- **Core Theorem.** *Let* $\mathcal{M}_0$ *be a model of* B *of any cardinality. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that admits a full satisfaction class.*
- **Proof:** Let F_0 be the set of *atomic* $\mathcal{N}$-formulas and let S_0 be the obvious satisfaction predicate for F_0. Then by the Lemma there is an elementary extension $\mathcal{M}_1$ of $\mathcal{M}_0$ that carries a $\mathrm{Form}_{\mathrm{B}}^{\mathcal{M}_0}$ satisfaction class. Thanks to Lemma 3.1, this argument can be carried out countably many times to yield two sequences $\langle \mathcal{M}_i : i \in \omega \rangle$ and $\langle S_i : i \in \omega \rangle$ that satisfy the following two properties:
- (1) $\mathcal{M}_i \prec \mathcal{M}_{i+1}$;

The Core Construction (7)

- **Core Theorem.** *Let* $\mathcal{M}_0$ *be a model of* B *of any cardinality. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that admits a full satisfaction class.*
- **Proof:** Let F_0 be the set of *atomic* $\mathcal{N}$-formulas and let S_0 be the obvious satisfaction predicate for F_0. Then by the Lemma there is an elementary extension $\mathcal{M}_1$ of $\mathcal{M}_0$ that carries a $\mathrm{Form}_{\mathrm{B}}^{\mathcal{M}_0}$ satisfaction class. Thanks to Lemma 3.1, this argument can be carried out countably many times to yield two sequences $\langle \mathcal{M}_i : i \in \omega \rangle$ and $\langle S_i : i \in \omega \rangle$ that satisfy the following two properties:
- (1) $\mathcal{M}_i \prec \mathcal{M}_{i+1}$;
- (2) S_{i+1} is a $\mathrm{Form}^{\mathcal{M}_i}$-satisfaction class on $\mathcal{M}_{i+1}$ for each $i \in \omega$.

The Core Construction (7)

- **Core Theorem.** *Let $\mathcal{M}_0$ be a model of* B *of any cardinality. There is an elementary extension $\mathcal{M}$ of $\mathcal{M}_0$ that admits a full satisfaction class.*
- **Proof:** Let F_0 be the set of *atomic* $\mathcal{N}$-formulas and let S_0 be the obvious satisfaction predicate for F_0. Then by the Lemma there is an elementary extension $\mathcal{M}_1$ of $\mathcal{M}_0$ that carries a $\mathrm{Form}_{\mathrm{B}}^{\mathcal{M}_0}$ satisfaction class. Thanks to Lemma 3.1, this argument can be carried out countably many times to yield two sequences $\langle \mathcal{M}_i : i \in \omega \rangle$ and $\langle S_i : i \in \omega \rangle$ that satisfy the following two properties:
- (1) $\mathcal{M}_i \prec \mathcal{M}_{i+1}$;
- (2) S_{i+1} is a $\mathrm{Form}^{\mathcal{M}_i}$-satisfaction class on $\mathcal{M}_{i+1}$ for each $i \in \omega$.

- $$\mathcal{M} := \bigcup_{i\in\omega} \mathcal{M}_i, \text{ and } S := \bigcup_{i\in\omega} S_i.$$

The Core Construction (8)

The Core Construction (8)

- **Corollary.** B^{FS} *is a conservative extension of* B *for every base theory* B.

- **Corollary.** B^{FS} *is a conservative extension of* B *for every base theory* B**.**

- **Theorem.** *Let* $<_{\mathbb{L}}$ *be a* B*-definable linear order on* $\mathbb{N}$ in the sense of B. *Every model of* B *has an elementary extension to a model that expands to* $\mathrm{B}^{\mathsf{FS}}_{\mathbb{L}}$. *Consequently,* $\mathrm{B}^{\mathsf{FS}}_{\mathbb{L}}$ *is conservative over* B *for every base theory*.

The Arithmetization of the Core Construction (1)

- **Definition.**
  (a) Suppose $\mathcal{M}$ is a model of some base theory, and $\mathcal{N}$ is a structure in a finite language $\mathcal{L}$. $\mathcal{N}$ is *strongly interpretable* in $\mathcal{M}$ if $\mathcal{M}$ can interpret an isomorphic copy $\mathcal{N}_0$ of $\mathcal{N}$; and moreover there is an $\mathcal{M}$-definable F-satisfaction class S on $\mathcal{N}_0$, where F is the collection of all $\mathcal{L}$-formulas in the sense of $\mathcal{M}$.

- **Definition.**
  (a) Suppose $\mathcal{M}$ is a model of some base theory, and $\mathcal{N}$ is a structure in a finite language $\mathcal{L}$. $\mathcal{N}$ is *strongly interpretable* in $\mathcal{M}$ if $\mathcal{M}$ can interpret an isomorphic copy $\mathcal{N}_0$ of $\mathcal{N}$; and moreover there is an $\mathcal{M}$-definable F-satisfaction class S on $\mathcal{N}_0$, where F is the collection of all $\mathcal{L}$-formulas in the sense of $\mathcal{M}$.
- **(b)** B *strongly interprets* $\mathrm{B}_0^{\mathrm{FS}}$, *i.e., every model* $\mathcal{M} \models \mathrm{B}$ *strongly interprets a structure* $(\mathcal{N}, S) \models \mathrm{B}_0^{\mathrm{FS}}$ *in a uniform manner.*

The Arithmetization of the Core Construction (1)

- **Definition.**
  (a) Suppose $\mathcal{M}$ is a model of some base theory, and $\mathcal{N}$ is a structure in a finite language $\mathcal{L}$. $\mathcal{N}$ is *strongly interpretable* in $\mathcal{M}$ if $\mathcal{M}$ can interpret an isomorphic copy $\mathcal{N}_0$ of $\mathcal{N}$; and moreover there is an $\mathcal{M}$-definable F-satisfaction class S on $\mathcal{N}_0$, where F is the collection of all $\mathcal{L}$-formulas in the sense of $\mathcal{M}$.
- **(b)** B *strongly interprets* $\mathrm{B}_0^{\mathrm{FS}}$, *i.e., every model* $\mathcal{M} \models \mathrm{B}$ *strongly interprets a structure* $(\mathcal{N}, S) \models \mathrm{B}_0^{\mathrm{FS}}$ *in a uniform manner.*

- **Theorem.** *Suppose* B *is an* **inductive** *base theory such that* $\mathrm{B} \vdash \mathrm{Con}(\mathrm{B}_0)$, *where* B_0 *is some r.e. base theory*. **Then** B *strongly interprets* $\mathrm{B}_0^{\mathrm{FS}}$.

The Arithmetization of the Core Construction (2)

- **Corollary.** *If* B *is an* **inductive** *theory, then:*

  1. $\mathsf{B} \vdash \mathsf{Con}(\mathsf{B}_0^{\mathsf{FS}})$ *for every finitely axiomatized base theory* $\mathsf{B}_0 \subseteq \mathsf{B}$.

The Arithmetization of the Core Construction (2)

- **Corollary.** *If* B *is an* **inductive** *theory, then:*

  1. $\mathsf{B} \vdash \mathsf{Con}(\mathsf{B}_0^{\mathsf{FS}})$ *for every finitely axiomatized base theory* $\mathsf{B}_0 \subseteq \mathsf{B}$.

- **2.** B^{IS} and B^{FS} *are not finitely axiomatizable for* **inductive** *base theories* B.

The Arithmetization of the Core Construction (3)

The Arithmetization of the Core Construction (3)

- **Theorem.** *The following statement* $(*)$ *is provable within* WKL_0 :

  $(*)$ *Every consistent base theory* B *has a model* $\mathcal{M}$ *that carries a full satisfaction class* S *and which has the property that the Tarskian satisfaction relation of* $(\mathcal{M}, S)$ *is coded by some* $X \subseteq \omega$.

- **Theorem.** *The following statement* $(*)$ *is provable within* WKL_0 :

  $(*)$ *Every consistent base theory* B *has a model* $\mathcal{M}$ *that carries a full satisfaction class* S *and which has the property that the Tarskian satisfaction relation of* $(\mathcal{M}, S)$ *is coded by some* $X \subseteq \omega$.

- **Theorem.** PRA $\vdash$ "B^{FS} *is conservative over* B" *for every r.e. base theory* B.

Pathological Satisfaction Classes

- **Definition.** For any standard formula σ of $\mathcal{L}_{\mathrm{B}}$, and for each $a \in \mathbb{N}^{\mathcal{M}}$, where $\mathcal{M}$ is some prescribed model of B, the 'formula' σ_a is defined by internal recursion in $\mathcal{M}_0$ via $\sigma_0 := \sigma$; and $\sigma_{n+1} := \sigma_n \vee \sigma_n$.

Pathological Satisfaction Classes

- **Definition.** For any standard formula σ of $\mathcal{L}_{\mathrm{B}}$, and for each $a \in \mathbb{N}^{\mathcal{M}}$, where $\mathcal{M}$ is some prescribed model of B, the 'formula' σ_a is defined by internal recursion in $\mathcal{M}_0$ via $\sigma_0 := \sigma$; and $\sigma_{n+1} := \sigma_n \vee \sigma_n$.

- **Theorem.** *Let $\sigma := \exists v_0\,(v_0 = v_0)$ (or $\sigma =$ any other logically valid sentence), and $\mathcal{M}_0$ be a model of* B *of any cardinality. Then $\mathcal{M}_0$ has an elementary extension $\mathcal{M}$ that carries a full satisfaction class S such that*

$$\{a \in \mathbb{N}^{\mathcal{M}} : \sigma_a \text{ is } S\text{-valid}\} = \omega.$$

Desirable Satisfaction Classes (1)

Desirable Satisfaction Classes (1)

- **Theorem.** *Let* $\mathcal{M}_0 \models \mathrm{B}$, *where* B *is a base theory. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that carries full satisfaction classes* S_1, S_2, *and* S_3 *such that:*

Desirable Satisfaction Classes (1)

- **Theorem.** *Let* $\mathcal{M}_0 \models \mathrm{B}$, *where* B *is a base theory. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that carries full satisfaction classes* S_1, S_2, *and* S_3 *such that:*
- $(\mathbf{1})$: S_1 *is* **schematically correct**;

Desirable Satisfaction Classes (1)

- **Theorem.** *Let* $\mathcal{M}_0 \models \mathrm{B}$, *where* B *is a base theory. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that carries full satisfaction classes* S_1, S_2, *and* S_3 *such that:*
- **(1)** : S_1 *is* **schematically correct**;
- **(2)** : S_2 *is both* **existentially and disjunctively correct**; *and*

Desirable Satisfaction Classes (1)

- **Theorem.** *Let* $\mathcal{M}_0 \models \mathrm{B}$, *where* B *is a base theory. There is an elementary extension* $\mathcal{M}$ *of* $\mathcal{M}_0$ *that carries full satisfaction classes* S_1, S_2, *and* S_3 *such that:*
- **(1)** : S_1 *is* **schematically correct***;*
- **(2)** : S_2 *is both* **existentially and disjunctively correct***; and*
- **(3)** : S_3 *is both* **extensional and alphabetically correct***.*

Desirable Satisfaction Classes (2)

Desirable Satisfaction Classes (2)

- *Moreover, if* B *is an* **inductive** *base theory, then* $\mathcal{M}$ *carries a full satisfaction class* S_4 *such that:*

  (**4**) : S_4 *is* $\Sigma_{\mathrm{B},\infty}$-**correct**,

Desirable Satisfaction Classes (2)

- *Moreover, if* B *is an* **inductive** *base theory, then* $\mathcal{M}$ *carries a full satisfaction class* S_4 *such that:*

  (**4**) : S_4 *is* $\Sigma_{\mathrm{B},\infty}$-**correct**,

- *and there is a family* $\{S_{5,s} : s \in \mathbb{N}^{\mathcal{M}}\}$ *of full satisfaction classes on* $\mathcal{M}$ *such that for each* $s \in \mathbb{N}^{\mathcal{M}}$ *there is a cut* I *of* $\mathbb{N}^{\mathcal{M}}$ *with* $I \models \mathrm{PA}$ *with* $s \in I$ *such that:*

  ($\mathbf{5}_s$) : $S_{5,s}$ *is* I-**deductively correct**.

Interpretability Issues (1)

Interpretability Issues (1)

- Let ACA be the strengthening of ACA_0 with the full scheme of induction. It has been long known that ACA and $\mathsf{PA}^{\mathsf{FIS}}$ are 'proof-theoretically equivalent'. The result below provides a more precise relationship between the two theories.

- Let ACA be the strengthening of ACA_0 with the full scheme of induction. It has been long known that ACA and $\mathsf{PA}^{\mathsf{FIS}}$ are 'proof-theoretically equivalent'. The result below provides a more precise relationship between the two theories.

- **Theorem.** *There is a sentence σ in the language of* ACA_0 *such that* $\mathsf{PA}^{\mathsf{FIS}}$ *and* $\mathsf{ACA} + \sigma$ *are bi-interpretable.*

Interpretability Issues (1)

- Let ACA be the strengthening of ACA_0 with the full scheme of induction. It has been long known that ACA and $\mathsf{PA}^{\mathsf{FIS}}$ are 'proof-theoretically equivalent'. The result below provides a more precise relationship between the two theories.
- **Theorem.** *There is a sentence σ in the language of* ACA_0 *such that* $\mathsf{PA}^{\mathsf{FIS}}$ *and* $\mathsf{ACA} + \sigma$ *are bi-interpretable.*
- **Theorem.** B^{IS} *and* B^{FS} *are both interpretable in* B *for every* **inductive** *recursively axiomatizable base theory* B.

Interpretability Issues (2)

Interpretability Issues (2)

- **Theorem** (Interpretability among PA, $\mathsf{PA}^{\mathsf{IS}}$, $\mathsf{PA}^{\mathsf{FS}}$, and ACA_0).

Interpretability Issues (2)

- **Theorem** (Interpretability among PA, $\mathsf{PA}^{\mathsf{IS}}$, $\mathsf{PA}^{\mathsf{FS}}$, and ACA_0).
- **(a)** *The theories* $\{\mathsf{PA}, \mathsf{PA}^{\mathsf{IS}}, \mathsf{PA}^{\mathsf{FS}}\}$ *are mutually interpretable.*

Interpretability Issues (2)

- **Theorem** (Interpretability among PA, $\mathrm{PA}^{\mathrm{IS}}$, $\mathrm{PA}^{\mathrm{FS}}$, and ACA_0).
- **(a)** *The theories* $\{\mathrm{PA},\ \mathrm{PA}^{\mathrm{IS}},\ \mathrm{PA}^{\mathrm{FS}}\}$ *are mutually interpretable.*
- **(b)** *Each of the theories* $\{\mathrm{PA},\ \mathrm{PA}^{\mathrm{IS}},\ \mathrm{PA}^{\mathrm{FS}}\}$ *is interpretable in* ACA_0, *but none of them interprets* ACA_0.
- **(c)** *No pair of the theories* $\{\mathrm{PA},\ \mathrm{PA}^{\mathrm{FS}},\ \mathrm{PA}^{\mathrm{IS}},\mathrm{ACA}_0\}$ *are bi-interpretable.*

Interpretability Issues (3)

Interpretability Issues (3)

- **Theorem.** *If* B *is a consistent* **finitely axiomatizable** *base theory, then neither* B^{IS} *nor* B^{FS} *is interpretable in* B.