Tarskian and Kripkean Truth Definitions

Lavinia Picollo
Assistant Professor

Seminar: *Truth and the Truth Predicate*
Munich Center for Mathematical Philosophy
LMU Munich - Winter Semester 2017/18

1 Technical Preliminaries

1.1 First-Order Languages

A first-order language $\mathscr{L}$ is a set of formulae, that is, strings of symbols from a certain *vocabulary* $V_{\mathscr{L}}$, combined in a particular way, to be specified.

1.1.1 Vocabulary of $\mathscr{L}$

The vocabulary $V_{\mathscr{L}}$ of $\mathscr{L}$ is a *finite* set containing

- Individual variables: $x_1, x_2, \ldots, x_n, \ldots$
- Logical symbols: $\neg, \rightarrow, \forall, =$
- Auxiliary symbols: (,), ,

plus some of the following:

- Individual constants: $c_1, c_2, \ldots, c_n, \ldots$
- Function symbols: $f_1^1, f_2^1, \ldots, f_n^1, \ldots, f_1^2, f_2^2, \ldots, f_n^2, \ldots, f_1^m, f_2^m, \ldots, f_n^m, \ldots$
- Relation symbols: $=, R_1^1, R_2^1, \ldots, R_n^1, \ldots, R_1^2, R_2^2, \ldots, R_n^2, \ldots, R_1^m, R_2^m, \ldots, R_n^m, \ldots$

1.1.2 Terms of $\mathscr{L}$

We define the class of *terms of* $\mathscr{L}$ recursively as follows:

- Individual constants in $V_{\mathscr{L}}$ are terms of $\mathscr{L}$.
- Individual variables are terms of $\mathscr{L}$.
- If $t_1, \ldots, t_n$ are terms of $\mathscr{L}$ and $f_i^n \in V_{\mathscr{L}}$, then $f_i^n(t_1, \ldots, t_n)$ is a term of $\mathscr{L}$.
- Only strings of symbols obtained applying the previous clauses a finite number of times are terms of $\mathscr{L}$.

Closed terms are those in which no variables occur.

1.1.3 Formulae of $\mathscr{L}$

We define the class of *formulae (wff) of* $\mathscr{L}$ recursively as follows:

- If s, t are terms of $\mathscr{L}$, then $s = t$ is a(n atomic) wff of $\mathscr{L}$.
- If $t_1, \ldots, t_n$ are terms of $\mathscr{L}$ and $R_i^n \in V_{\mathscr{L}}$, then $R_i^n(t_1, \ldots, t_n)$ is a(n atomic) wff of $\mathscr{L}$.
- If φ is a wff of $\mathscr{L}$, then $\neg\varphi$ is a wff of $\mathscr{L}$.
- If φ and ψ are wff of $\mathscr{L}$, then $(\varphi \wedge \psi)$ is a wff of $\mathscr{L}$.
- If φ is a wff of $\mathscr{L}$ and v an individual variable, then $\forall v\varphi$ is a wff of $\mathscr{L}$.
- Only strings of symbols obtained the previous clauses a finite number of times are wff of $\mathscr{L}$.

Sentences of $\mathscr{L}$ are wff in which every individual variable v that occurs in them occurs under the scope of $\forall v$. That is, there are no free (unbound) variables in sentences.

The following formulae on the left abbreviate the respective formulae on the right:

- $s \neq t := \neg s = t$;
- $\varphi \wedge \psi := \neg(\varphi \to \neg\psi)$;
- $\varphi \vee \psi := \neg\varphi \to \psi$;
- $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi)$;
- $\exists v\varphi := \neg\forall v\neg\varphi$;
- $\exists! v\varphi(v) := \exists v(\varphi(v) \wedge \forall u(\varphi(u) \to u = v))$.

Let $\mathscr{L}\restriction$ result form $\mathscr{L}$ by removing all formulae containing variables, free or bound.

1.1.4 The Language of Peano Arithmetic

$\mathscr{L}_{\mathsf{PA}}$'s non-logical symbols are the individual constant 0, the unary function symbol S, and the binary function symbols $+$ and $\times$.

Let $\mathscr{L}$ be any first-order language. $\mathscr{L}_{\mathsf{PA}}$ contains a name $\ulcorner\epsilon\urcorner$ for each expression ϵ of $\mathscr{L}$, via an effective and monotonic coding of expressions of $\mathscr{L}$ with numbers.

1.1.5 The Language of Real Analysis

$\mathscr{L}_{\mathsf{RA}}$'s non-logical symbols are the individual constants 0 and 1, the binary function symbols $+$ and $\times$, and the binary relation symbol $<$.

1.1.6 The Language of Zermelo-Fraenkel Set Theory

$\mathscr{L}_{\mathsf{ZF}}$'s non-logical symbols are the binary relation symbol $\in$ (primitive) and the individual constant $\varnothing$. $\mathscr{L}_{\mathsf{ZF}}$ will be the *metalanguage* for all the truth definitions introduced here.

$\mathscr{L}_{\mathsf{ZF}}$ also contains a name $\ulcorner\epsilon\urcorner$ for each expression ϵ of every first-order language $\mathscr{L}$ via an effective and monotonic coding of expressions of $\mathscr{L}$ with sets. Additionally, $\mathscr{L}_{\mathsf{ZF}}$ has formulae (on the left) defining the following properties and functions (on the right):[1]

$$\begin{aligned}
Var(v) &:= v \text{ is an individual variable}\\
Term_{\mathscr{L}}(v) &:= v \text{ is a term of } \mathscr{L}\\
CTerm_{\mathscr{L}}(v) &:= v \text{ is a closed term of } \mathscr{L}\\
Form_{\mathscr{L}}(v) &:= v \text{ is a wff of } \mathscr{L}\\
Sent_{\mathscr{L}}(v) &:= v \text{ is a sentence of } \mathscr{L}\\
u \frown v &:= \text{the concatenation of expressions } u \text{ and } v\\
s(u,v,w) &:= \text{the result of substituting all free occurrences of variable } v \text{ in formula } u\\
&\quad\ \text{with term } w, \text{ if } w \text{ is free for } v \text{ in } u.
\end{aligned}$$

In $\mathscr{L}_{\mathsf{ZF}}$ we can also define what it means to be relation, a function (both set-theoretic entities), and many specific relations and functions.

1.2 Proof-Theory for First-Order Languages

1.2.1 A Calculus for First-Order Logic

We introduce the Hilbert-style calculus $\mathsf{FOL}^=$. Let φ, ψ, χ be formulae of a first-order language, t a term, and v a variable.

The *axioms of* $\mathsf{FOL}^=$ are given by all instances of the following schemata:

A1 $\varphi \to (\psi \to \varphi)$

A2 $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$

A3 $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$

A4 $\forall v\varphi \to \varphi[t/v]$, if t is free for v in φ.

=1 $\forall x\ x = x$

=2 $\forall x \forall y (x = y \to (\varphi(x) \to \varphi(y)))$, if y is free for x in φ.

$\mathsf{FOL}^=$ also consists of the following *rules of inference*:

$$\frac{\varphi, \varphi \to \psi}{\psi}\ (\textit{Modus Ponens}) \qquad\qquad \frac{\varphi \to \psi}{\varphi \to \forall v\psi}\ (\textit{Generalisation})$$

if v is not free in φ of in a premise.

1.2.2 Proof-Theoretic Validity in $\mathsf{FOL}^=$

Let $\mathscr{L}$ be a first-order language, $\varphi \in \mathscr{L}$, and $\Gamma \subseteq \mathscr{L}$. φ is a *proof-theoretic consequence* of Γ in $\mathsf{FOL}^=$ ($\Gamma \vdash \varphi$) iff there is a sequence of wff of $\mathscr{L}$ s.t. each member is either an axiom of $\mathsf{FOL}^=$, or a member of Γ, or was obtained by one of the rules of inference of $\mathsf{FOL}^=$ from previous members of the sequence, and the last one is φ.

If Γ is empty, φ is a *logical theorem* ($\vdash \varphi$).

[1] Note that, given any two expressions ϵ_1 and ϵ_2 of $\mathscr{L}$, $\ulcorner\epsilon_1\urcorner \frown \ulcorner\epsilon_2\urcorner = \ulcorner\epsilon_1\epsilon_2\urcorner$.

1.2.3 Theories in $\mathsf{FOL}^=$

A *theory* in $\mathsf{FOL}^=$ is a set of formulae of a first-order language $\mathscr{L}$, closed under proof-theoretic consequence in $\mathsf{FOL}^=$ (i.e. if φ is a proof-theoretic consequence of some members of the theory, φ is also a member of the theory).

1.2.4 Peano Arithmetic

$\mathsf{PA} \subseteq \mathscr{L}_{\mathsf{PA}}$ in $\mathsf{FOL}^=$ consists of the following axioms:

PA1 $\forall x(Sx \neq 0)$

PA2 $\forall x \forall y(Sx = Sy \to x = y)$

PA3 $\forall x(x + 0 = x)$

PA4 $\forall x \forall y(x + Sy = S(x + y))$

PA5 $\forall x(x \times 0 = 0)$

PA6 $\forall x \forall y(x \times Sy = x \times y + x)$

Induction schema. $\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(Sx)) \to \forall x \varphi(x)$

Theorem 1 (Diagonalisation)**.** *For every $\varphi(x) \in \mathscr{L}_{\mathsf{PA}}$ there is a $\psi \in \mathscr{L}_{\mathsf{PA}}$ s.t.*

$$\mathsf{PA} \vdash \varphi(\ulcorner \psi \urcorner) \leftrightarrow \psi$$

1.2.5 Real Analysis

$\mathsf{RA} \subseteq \mathscr{L}_{\mathsf{RA}}$ in $\mathsf{FOL}^=$ consists of the following axioms:

RA1 $\forall x \forall y \forall z((x + y) + z = x + (y + z))$

RA2 $\forall x \forall y(x + y = y + x)$

RA3 $\exists x \forall y(x + y = x)$

Def. of 0**.** $\forall x(x = 0 \leftrightarrow \forall y(y + x = y))$

RA4 $\forall x \exists y(x + y = 0)$

RA5 $\forall x \forall y \forall z((x \times y) \times z = x \times (y \times z))$

RA6 $\forall x \forall y(x \times y = y \times x)$

RA7 $\exists x \forall y(y \times x = y)$

Def. of 1**.** $\forall x(x = 1 \leftrightarrow \forall y(y \times x = y))$

RA8 $\forall x(x \neq 0 \to \exists y(x \times y = 1))$

RA9 $\forall x \forall y \forall z((x + y) \times z = (x \times z) + (y \times z))$

RA10 $\forall x(x < 0 \vee x = 0 \vee 0 < x))$

RA11 $\forall x \forall y(0 < x \wedge 0 < y \to 0 < x + y \wedge 0 < x \times y)$

RA12 $\forall x \forall y \forall z(x < y \to x + y < y + z)$

Completeness schema. $\exists x \varphi(x) \wedge \exists x \forall y(\varphi(y) \to y < x \vee y = x) \to \exists x(\forall y(\varphi(y) \to y < x \vee y = x) \wedge \forall z(\forall y(\varphi(y) \to y < z \vee y = z) \to x < z))$

1.2.6 Zermelo-Fraenkel Set Theory

$\mathsf{ZF} \subseteq \mathscr{L}_{\mathsf{ZF}}$ in $\mathsf{FOL}^=$ consists of the following axioms and definitons:

Separation schema. $\forall x_1 \ldots \forall x_n \forall y \exists z \forall w(w \in z \leftrightarrow w \in y \land \varphi(w, x_1, \ldots, x_n))$

Extensionality. $\forall x \forall y(\forall z(z \in x \leftrightarrow z \in y) \to x = y)$

Empty set. $\exists x \forall y \neg y \in x$

Def. of $\varnothing$. $\forall x(x = \varnothing \leftrightarrow \forall y \neg y \in x)$

Pairing. $\forall x \forall y \exists z \forall w(w \in z \leftrightarrow w = x \lor w = y)$

Union. $\forall x \exists y \forall z(z \in y \leftrightarrow \exists w(w \in x \land z \in w))$

Power set. $\forall x \exists y \forall z(z \in y \leftrightarrow \forall w(w \in z \to w \in x))$

Infinity. $\exists x(\varnothing \in x \land \forall y(y \in x \to \{y\} \in x))$

Replacement schema. $\forall x(\forall y(y \in x \to \exists! z\varphi(y, z)) \to \exists y \forall z(z \in y \leftrightarrow \exists w(w \in x \land \varphi(w, z))))$

Foundation. $\forall x(x \neq \varnothing \to \exists y(y \in x \land \neg \exists z(z \in y \land z \in x)))$

Induction schema. $\forall x(\forall y(y \in x \to \varphi(y)) \to \varphi(x)) \to \forall x \varphi(x)$

where $\varphi(\{t_1, \ldots, t_n\})$ abbreviates $\varphi(v) \land \forall u(u \in v \leftrightarrow u = t_1 \lor \cdots \lor u = t_n)$.

1.3 Tarski's Undefinability Theorem

Theorem 2. *Let $\mathscr{L}$ be a first-order language containing a term $\ulcorner \varphi \urcorner$ denoting each sentence $\varphi \in \mathscr{L}$, and let $\mathsf{Th} \subseteq \mathscr{L}$ be s.t., for every $\varphi(x) \in \mathscr{L}$, there is a $\psi \in \mathscr{L}$ s.t. $\mathsf{Th} \vdash \varphi(\ulcorner \psi \urcorner) \leftrightarrow \psi$. If Th is consistent, there is no $\tau(x) \in \mathscr{L}$ s.t., for every φ,*

$$\mathsf{Th} \vdash \tau(\ulcorner \varphi \urcorner) \leftrightarrow \varphi.$$

No theory in $\mathsf{FOL}^=$ formulated in a language that contains names for its own expressions, in which a version of diagonalisation holds, can have a formula expressing its truth for its own formulae. As a consequence, a truth definition for a language, the *object language*, must always be carried in a theory, the *metatheory*, formulated in different language, the *metalanguage*. The metalanguage must be *essentially richer* than the object language.

2 Tarskian Truth

In this section we provide increasingly complex examples of Tarskian definitions of truth for several first-order languages, our object languages, in expansions of ZF, our metalanguages. These definitions are all of a recursive nature.

2.1 Truth for a restricted set-theoretic language

We define truth for $\mathscr{L}_{\mathsf{ZF}}\restriction$ in ZF, as follows:

$$\begin{aligned} T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(x) \leftrightarrow_{def} \; & ((x = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing) \lor (x = \ulcorner \varnothing \in \varnothing \urcorner \land \varnothing \in \varnothing) \lor \\ & \exists y (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land x = \ulcorner \neg \urcorner \frown y \land \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y)) \lor \\ & \exists y \exists z (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z) \land x = \ulcorner (\urcorner \frown y \frown \ulcorner \land \urcorner \frown z \frown \ulcorner) \urcorner \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z))) \end{aligned}$$

Proposition 3. *The following are theorems of* ZF*:*

- $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$*, for every sentence* $\varphi \in \mathscr{L}_{\mathsf{ZF}}\restriction$
- $\forall x (T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(x) \to Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(x))$

Proof. The proof of the first claim is by induction on the complexity of φ. If φ is atomic, then it is of the form $\varnothing = \varnothing$ or $\varnothing \in \varnothing$. For the first case, we reason as follows:

1. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varnothing = \varnothing \urcorner)$ — assumption
2. $(\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing) \lor (\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing \in \varnothing \urcorner \land \varnothing \in \varnothing) \lor$
$\exists y (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land \ulcorner \varnothing = \varnothing \urcorner = \ulcorner \neg \urcorner \frown y \land \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y)) \lor$
$\exists y \exists z (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z) \land \ulcorner \varnothing = \varnothing \urcorner = \ulcorner (\urcorner \frown y \frown \ulcorner \land \urcorner \frown z \frown \ulcorner) \urcorner \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z))$ Def. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}$, 1
3. $\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing$ — ZF, 2
4. $\varnothing = \varnothing$ — 3
5. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varnothing = \varnothing \urcorner) \to \varnothing = \varnothing$ — 1-4

and

1. $\varnothing = \varnothing$ — assumption
2. $\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing$ — =1, 1
3. $(\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing) \lor (\ulcorner \varnothing = \varnothing \urcorner = \ulcorner \varnothing \in \varnothing \urcorner \land \varnothing \in \varnothing) \lor$
$\exists y (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land \ulcorner \varnothing = \varnothing \urcorner = \ulcorner \neg \urcorner \frown y \land \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y)) \lor$
$\exists y \exists z (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z) \land \ulcorner \varnothing = \varnothing \urcorner = \ulcorner (\urcorner \frown y \frown \ulcorner \land \urcorner \frown z \frown \ulcorner) \urcorner \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z))$ — 2
4. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varnothing = \varnothing \urcorner)$ — Def. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}$, 3
5. $\varnothing = \varnothing \to T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varnothing = \varnothing \urcorner)$ — 1-4

The case for $\varnothing \in \varnothing$ is analogous. Try to prove it!

For the inductive step, assume $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$ for every sentence φ of complexity less than n (the *inductive hypothesis*), and let ψ be of complexity n. ψ is either a negation of a conjunction. If the former is the case, there is a sentence χ s.t. ψ is $\neg\chi$. Thus, χ is of complexity $n - 1$. Thus, by inductive hypothesis, we have that $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \chi \urcorner) \leftrightarrow \chi$. We then reason as follows:

1. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \psi \urcorner)$ — assumption
2. $(\ulcorner \psi \urcorner = \ulcorner \varnothing = \varnothing \urcorner \land \varnothing = \varnothing) \lor (\ulcorner \psi \urcorner = \ulcorner \varnothing \in \varnothing \urcorner \land \varnothing \in \varnothing) \lor$
$\exists y (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land \ulcorner \psi \urcorner = \ulcorner \neg \urcorner \frown y \land \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y)) \lor$
$\exists y \exists z (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z) \land \ulcorner \psi \urcorner = \ulcorner (\urcorner \frown y \frown \ulcorner \land \urcorner \frown z \frown \ulcorner) \urcorner \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z))$ Def. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}$, 1
3. $\exists y (Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \land \ulcorner \psi \urcorner = \ulcorner \neg \urcorner \frown y \land \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y))$ — ZF, 3
4. $\neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner \chi \urcorner)$ — ZF, 3

5. $\neg\chi$[2] inductive hypothesis
6. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner\psi\urcorner) \to \psi$ 1-5

and

1. ψ assumption
2. $\neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner\chi\urcorner)$ inductive hypothesis, 1
3. $\exists y(Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \wedge \ulcorner\psi\urcorner = \ulcorner\neg\urcorner\frown y \wedge \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y))$ ZF, 2
4. $(\ulcorner\psi\urcorner = \ulcorner\varnothing = \varnothing\urcorner \wedge \varnothing = \varnothing) \vee (\ulcorner\psi\urcorner = \ulcorner\varnothing \in \varnothing\urcorner \wedge \varnothing \in \varnothing) \vee$
   $\exists y(Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \wedge \ulcorner\psi\urcorner = \ulcorner\neg\urcorner\frown y \wedge \neg T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y)) \vee$
   $\exists y\exists z(Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \wedge Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z) \wedge \ulcorner\psi\urcorner = \ulcorner(\urcorner\frown y\frown\ulcorner\wedge\urcorner\frown z\frown\ulcorner)\urcorner \wedge T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(y) \wedge T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(z))$ 3
5. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner\psi\urcorner)$ Def. $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}$, 4
6. $\psi \to T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(\ulcorner\psi\urcorner)$ 1-5

The case for conjunction can be proved in a similar fashion. Try to prove it!

$\forall x(T_{\mathscr{L}_{\mathsf{ZF}}\restriction}(x) \to Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}(x))$ follows directly from the definition of $T_{\mathscr{L}_{\mathsf{ZF}}\restriction}$ and that of $Sent_{\mathscr{L}_{\mathsf{ZF}}\restriction}$. $\square$

2.2 Truth for a very simple language

Let $\mathscr{L}_S$ be a first-order language whose non-logical symbols are two individual constants n and m, intended to denote Norbert and Mitch, resp., a monadic predicate P intended to express the property of being a philosopher, and a binary relation symbol H intended to express the relation of begin the human of. Let $\mathscr{L}^S_{\mathsf{ZF}}$ extend $\mathscr{L}_{\mathsf{ZF}}$ with an individual constant a to denote Norbert, b to denote Mitch, p for the set of philosophers and h for the set of order pairs s.t. the first component is the human of the second one. We define truth for $\mathscr{L}_S\restriction$ in ZF_S, an extension of ZF with adequate definitions/axioms for the new vocabulary, as follows:

$$\begin{aligned}T_{\mathscr{L}_S\restriction}(x) \leftrightarrow_{def}\ & (x = \ulcorner n = n\urcorner \wedge a = a) \vee (x = \ulcorner n = m\urcorner \wedge a = b) \vee (x = \ulcorner m = n\urcorner \wedge b = a) \vee \\ & (x = \ulcorner m = m\urcorner \wedge b = b) \vee (x = \ulcorner Pn\urcorner \wedge a \in p) \vee (x = \ulcorner Pm\urcorner \wedge b \in p) \vee \\ & (x = \ulcorner Hnn\urcorner \wedge \langle a, a\rangle \in h) \vee (x = \ulcorner Hnm\urcorner \wedge \langle a, b\rangle \in h) \vee \\ & (x = \ulcorner Hmn\urcorner \wedge \langle b, a\rangle \in h) \vee (x = \ulcorner Hmm\urcorner \wedge \langle b, b\rangle \in h) \vee \\ & \exists y(Sent_{\mathscr{L}_S\restriction}(y) \wedge x = \ulcorner\neg\urcorner\frown y \wedge \neg T_{\mathscr{L}_S\restriction}(y)) \vee \\ & \exists y\exists z(Sent_{\mathscr{L}_S\restriction}(y) \wedge Sent_{\mathscr{L}_S\restriction}(z) \wedge x = \ulcorner(\urcorner\frown y\frown\ulcorner\wedge\urcorner\frown z\frown\ulcorner)\urcorner \wedge T_{\mathscr{L}_S\restriction}(y) \wedge T_{\mathscr{L}_S\restriction}(z))\end{aligned}$$

Proposition 4. *The following are theorems of* ZF_S*:*

- $T_{\mathscr{L}_S\restriction}(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$*, for every sentence* $\varphi \in \mathscr{L}_S\restriction$
- $\forall x(T_{\mathscr{L}_S\restriction}(x) \to Sent_{\mathscr{L}_S\restriction}(x))$

2.3 Truth for a restricted arithmetical language

Since $\mathscr{L}_{\mathsf{PA}}$ contains complex terms, we need to take a detour. We extend $\mathscr{L}_{\mathsf{ZF}}$ with '0', to denote 0, 'f_S', for the function that maps each natural number to it's successor, 'f_+' for the

[2] Note that ψ *is* $\neg\chi$.

one that maps every pair of natural numbers to their sum, and '$f_\times$' for the one that maps them to their product. Accordingly, we recursively define, in an extension $\mathsf{ZF_{PA}}$ of ZF with the usual definitions/axioms for the new symbols, a function val that assigns to each closed term of $\mathscr{L}_{\mathsf{PA}}$ a value, as follows:

- $val(\ulcorner 0 \urcorner) = 0$;
- $\forall x(CTerm_{\mathscr{L}_{\mathsf{PA}}}(x) \rightarrow val(\ulcorner S \urcorner \frown x) = f_S(val(x)))$;
- $\forall x \forall y(CTerm_{\mathscr{L}_{\mathsf{PA}}}(x) \wedge CTerm_{\mathscr{L}_{\mathsf{PA}}}(y) \rightarrow val(x \frown \ulcorner + \urcorner \frown y) = f_+(val(x), val(y)))$;
- $\forall x \forall y(CTerm_{\mathscr{L}_{\mathsf{PA}}}(x) \wedge CTerm_{\mathscr{L}_{\mathsf{PA}}}(y) \rightarrow val(x \frown \ulcorner \times \urcorner \frown y) = f_\times(val(x), val(y)))$.

Now we are in a position to define truth for $\mathscr{L}_{\mathsf{PA}}\!\upharpoonright$ in $\mathsf{ZF_{PA}}$:

$$\begin{aligned} T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(x) \leftrightarrow_{def} \ & \exists y \exists z((Cterm_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(y) \wedge Cterm_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(z) \wedge x = y \frown \ulcorner = \urcorner \frown z \wedge val(y) = val(z)) \vee \\ & (Sent_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(y) \wedge x = \ulcorner \neg \urcorner \frown y \wedge \neg T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(y)) \vee \\ & (Sent_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(y) \wedge Sent_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(z) \wedge x = \ulcorner (\urcorner \frown y \frown \ulcorner \wedge \urcorner \frown z \frown \ulcorner) \urcorner \wedge T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(y) \wedge T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(z))) \end{aligned}$$

Proposition 5. *The following are theorems of* $\mathsf{ZF_{PA}}$*:*

- $T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$, *for every sentence* $\varphi \in \mathscr{L}_{\mathsf{PA}}\!\upharpoonright$
- $\forall x(T_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(x) \rightarrow Sent_{\mathscr{L}_{\mathsf{PA}}\upharpoonright}(x))$

2.4 Truth for the language of arithmetic

We define truth for $\mathscr{L}_{\mathsf{PA}}$ in $\mathsf{ZF_{PA}}$ as follows:

$$\begin{aligned} T_{\mathscr{L}_{\mathsf{PA}}}(x) \leftrightarrow_{def} \ & \exists y \exists z((CTerm_{\mathscr{L}_{\mathsf{PA}}}(y) \wedge CTerm_{\mathscr{L}_{\mathsf{PA}}}(z) \wedge x = y \frown \ulcorner = \urcorner \frown z \wedge val(x) = val(t)) \vee \\ & (Sent_{\mathscr{L}_{\mathsf{PA}}}(y) \wedge x = \ulcorner \neg \urcorner \frown y \wedge \neg T_{\mathscr{L}_{\mathsf{PA}}}(y)) \vee \\ & (Sent_{\mathscr{L}_{\mathsf{PA}}}(y) \wedge Sent_{\mathscr{L}_{\mathsf{PA}}}(z) \wedge x = \ulcorner (\urcorner \frown y \frown \ulcorner \wedge \urcorner \frown z \frown \ulcorner) \urcorner \wedge T_{\mathscr{L}_{\mathsf{PA}}}(y) \wedge T_{\mathscr{L}_{\mathsf{PA}}}(z)) \vee \\ & (Var(y) \wedge Sent_{\mathscr{L}_{\mathsf{PA}}}(s(z, \ulcorner 0 \urcorner, y)) \wedge x = \ulcorner \forall \urcorner \frown y \frown z \wedge \\ & \qquad \forall w(Cterm_{\mathscr{L}_{\mathsf{PA}}}(w) \rightarrow T_{\mathscr{L}_{\mathsf{PA}}}(s(z, w, y))))) \end{aligned}$$

Proposition 6. *The following are theorems of* $\mathsf{ZF_{PA}}$*:*

- $T_{\mathscr{L}_{\mathsf{PA}}}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$, *for every sentence* $\varphi \in \mathscr{L}_{\mathsf{PA}}$
- $\forall x(T_{\mathscr{L}_{\mathsf{PA}}}(x) \rightarrow Sent_{\mathscr{L}_{\mathsf{PA}}}(x))$

2.5 Truth for the language of real anaysis

Since the quantifiers in $\mathscr{L}_{\mathsf{RA}}$ are intended to range over non-denumerably many objects (the natural numbers), using the substitution function s will not do. We need to take a detour via assignments and satisfaction.

Expand $\mathscr{L}_{\mathsf{ZF}}$ with '$\mathbb{R}$', to denote the set of real numbers, '0', to denote 0 ($\in \mathbb{R}$), '1', to denote 1 ($\in \mathbb{R}$), 'f_+', for the function that maps every pair of real numbers to their sum, '$f_\times$' for the one that maps them to their product, and '$\prec$' for the set of ordered pairs of real numbers such that the first one is smaller than the second one. Let $\mathsf{ZF_{RA}}$ extend ZF with definitions/axioms for the new symbols.

An assignment for $\mathscr{L}_{\mathsf{RA}}$ is a function $\sigma : \{x : Term_{\mathscr{L}_{\mathsf{RA}}}(x)\} \to \mathbb{R}$ (it assigns members of $\mathbb{R}$ to each term of $\mathscr{L}_{\mathsf{RA}}$) s.t.

- $\forall x(Var(x) \to \sigma(x) \in \mathbb{R})$; (this is redundant)
- $\sigma(\ulcorner 0 \urcorner) = 0 \wedge \sigma(\ulcorner 1 \urcorner) = 1$;
- $\forall x \forall y(Term_{\mathscr{L}_{\mathsf{RA}}}(x) \wedge Term_{\mathscr{L}_{\mathsf{RA}}}(y) \to \sigma(x \frown \ulcorner + \urcorner \frown y) = f_+(\sigma(x), \sigma(y)))$;
- $\forall x \forall y(Term_{\mathscr{L}_{\mathsf{RA}}}(x) \wedge Term_{\mathscr{L}_{\mathsf{RA}}}(y) \to \sigma(x \frown \ulcorner \times \urcorner \frown y) = f_\times(\sigma(x), \sigma(y)))$.

The property of being an assignment for $\mathscr{L}_{\mathsf{RA}}$ is definable in $\mathsf{ZF}_{\mathsf{RA}}$, say, by the formula $A_{\mathscr{L}_{\mathsf{RA}}}(v)$. We define satisfaction of a *formula* of $\mathscr{L}_{\mathsf{RA}}$ by an assignment in $\mathsf{ZF}_{\mathsf{RA}}$ as follows:

$$
\begin{aligned}
Sat_{\mathscr{L}_{\mathsf{RA}}}(x, \sigma) \leftrightarrow_{def} \; & Form_{\mathscr{L}_{\mathsf{RA}}}(x) \wedge A_{\mathscr{L}_{\mathsf{RA}}}(\sigma) \wedge \\
& \exists y \exists z((x = y \frown \ulcorner = \urcorner \frown z \wedge \sigma(y) = \sigma(z)) \vee \\
& \quad (x = y \frown \ulcorner < \urcorner \frown z \wedge \sigma(y) \prec \sigma(z)) \vee \\
& \quad (x = \ulcorner \neg \urcorner \frown y \wedge \neg Sat_{\mathscr{L}_{\mathsf{RA}}}(y, \sigma)) \vee \\
& \quad (x = \ulcorner (\urcorner \frown y \frown \ulcorner \wedge \urcorner \frown z \frown \ulcorner) \urcorner \wedge Sat_{\mathscr{L}_{\mathsf{RA}}}(y, \sigma) \wedge Sat_{\mathscr{L}_{\mathsf{RA}}}(z, \sigma)) \vee \\
& \quad (x = \ulcorner \forall \urcorner \frown y \frown z \wedge \\
& \quad \forall \sigma' \forall w((A_{\mathscr{L}_{\mathsf{RA}}}(\sigma') \wedge Var(w) \wedge \sigma(w) \neq \sigma'(w) \to w = y) \to Sat_{\mathscr{L}_{\mathsf{RA}}}(z, \sigma'))))
\end{aligned}
$$

Now we are in a position to define truth of a *sentence* of $\mathscr{L}_{\mathsf{RA}}$ in $\mathsf{ZF}_{\mathsf{RA}}$:

$$T_{\mathscr{L}_{\mathsf{RA}}}(x) \leftrightarrow_{def} Sent_{\mathscr{L}_{\mathsf{RA}}}(x) \wedge \forall \sigma(A_{\mathscr{L}_{\mathsf{RA}}}(\sigma) \to Sat_{\mathscr{L}_{\mathsf{RA}}}(x, \sigma))$$

Proposition 7. *The following are theorems of* $\mathsf{ZF}_{\mathsf{RA}}$*:*

- $T_{\mathscr{L}_{\mathsf{RA}}}(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$*, for every sentence* $\varphi \in \mathscr{L}_{\mathsf{RA}}$
- $\forall x(T_{\mathscr{L}_{\mathsf{RA}}}(x) \to Sent_{\mathscr{L}_{\mathsf{RA}}}(x))$

2.6 Truth in a model for arbitrary languages

Let $\mathscr{L}$ be a first-order language. A *model* of $\mathscr{L}$ is an ordered pair $\mathcal{M} = \langle d, \mathcal{I} \rangle$, in which d is a non-empty set, the *domain* of the model, and $\mathcal{I}$ is an *interpretation* function. It assigns an interpretation to every individual constant, function symbol, and relation symbol of $\mathscr{L}$ as follows:

- if c is an individual constant of $\mathscr{L}$, $\mathcal{I}(\ulcorner c \urcorner) \in d$;
- if f is an n-ary function symbol of $\mathscr{L}$, $\mathcal{I}(\ulcorner f \urcorner)$ is a function from d^n to d;
- if R is an n-ary relation symbol of $\mathscr{L}$, $\mathcal{I}(\ulcorner R \urcorner) \subseteq d^n$.

We can easily define models in $\mathscr{L}_{\mathsf{ZF}}$, say, by the formula $Model(v)$. Models are set-theoretic structures.

A *variable assignment* σ in $\mathcal{M} = \langle d, \mathcal{I} \rangle$ is a function $\sigma : \{x : Term_{\mathscr{L}}(x)\} \to d$ s.t.

- if c is an individual constant, $\sigma(c) = \mathcal{I}(c)$;
- if $t_1, \ldots, t_n$ are terms of $\mathscr{L}$, $\sigma(f_i^n(t_1, \ldots, t_n)) = \mathcal{I}(f_i^n)(\sigma(t_1), \ldots, \sigma(t_n))$.

The property of being an interpretation function assignment in a model $\mathcal{M}$ of $\mathscr{L}$ is definable in $\mathscr{L}_{\mathsf{ZF}}$, say, by the formula $A_{\mathscr{L}}(v, u)$. We define satisfaction of a formula of $\mathscr{L}$ in a model by an assignment in ZF as follows:

$$
\begin{aligned}
Sat_{\mathscr{L}}(x,\mathcal{M},\sigma) \leftrightarrow_{def} & Form_{\mathscr{L}}(x) \wedge Model(\mathcal{M}) \wedge A_{\mathscr{L}}(\mathcal{M},\sigma) \wedge \\
& \ldots \vee\ \exists y_1 \ldots \exists y_n (x = \ulcorner R_i^n \urcorner \frown y_1 \frown \ldots \frown y_n \wedge \langle \sigma(y_1) \ldots, \sigma(y_n) \rangle \in \mathcal{I}(\ulcorner R_i^n \urcorner)) \vee \ldots \\
& \exists y \exists z ((x = y \frown \ulcorner = \urcorner \frown z \wedge \sigma(y) = \sigma(z)) \vee \\
& (x = \ulcorner \neg \urcorner \frown y \wedge \neg Sat_{\mathscr{L}}(y,\mathcal{M},\sigma)) \vee \\
& (x = \ulcorner (\urcorner \frown y \frown \ulcorner \wedge \urcorner \frown z \frown \ulcorner) \urcorner \wedge Sat_{\mathscr{L}}(y,\mathcal{M},\sigma) \wedge Sat_{\mathscr{L}}(z,\mathcal{M},\sigma)) \vee \\
& (x = \ulcorner \forall \urcorner \frown y \frown z \wedge \\
& \ \forall \sigma' \forall w ((A_{\mathscr{L}}(\mathcal{M},\sigma') \wedge Var(w) \wedge \sigma(w) \neq \sigma'(w) \rightarrow w = y) \rightarrow Sat_{\mathscr{L}}(z,\mathcal{M},\sigma'))))
\end{aligned}
$$

The definition contains a disjunct

$$\exists y_1 \ldots \exists y_n (x = \ulcorner R_i^n \urcorner \frown y_1 \frown \ldots \frown y_n \wedge \langle \sigma(y_1) \ldots, \sigma(y_n) \rangle \in \mathcal{I}(\ulcorner R_i^n \urcorner))$$

for each n-ary relation symbol R_i^n of $\mathscr{L}$. (Recall $\mathscr{L}$ always contains finitely many relation symbols.)

We now define truth-in-$\mathcal{M}$ of a *sentence* of $\mathscr{L}$ directly in ZF, as follows:

$$T_{\mathscr{L}}(x,\mathcal{M}) \leftrightarrow_{def} Sent_{\mathscr{L}}(x) \wedge Model(\mathcal{M}) \wedge \forall \sigma (A_{\mathscr{L}}(\mathcal{M},\sigma) \rightarrow Sat_{\mathscr{L}}(x,\mathcal{M},\sigma))$$

The following are (more familiar, sloppy) notational variants of the definitions of satisfaction and truth in $\mathcal{M}$ just given:

$$
\begin{aligned}
\mathcal{M},\sigma \vDash \varphi \leftrightarrow_{def} & \ldots \vee (\varphi := R_i^n t_1 \ldots t_n \wedge \langle \sigma(t_1) \ldots, \sigma(t_n) \rangle \in \mathcal{I}(R_i^n)) \vee \ldots \\
& (\varphi := (s = t) \wedge \sigma(s) = \sigma(t)) \vee \\
& (\varphi =: \neg \psi \wedge \mathcal{M},\sigma \nvDash \psi) \vee \\
& (\varphi := (\psi \wedge \chi) \wedge \mathcal{M},\sigma \vDash \psi \wedge \mathcal{M},\sigma \vDash \chi) \vee \\
& (\varphi := \forall v \psi \wedge \forall \sigma' \forall u ((Var(u) \wedge \sigma(u) \neq \sigma'(u) \rightarrow u = v) \rightarrow \mathcal{M},\sigma' \vDash \psi))
\end{aligned}
$$

$$\mathcal{M} \vDash \varphi \leftrightarrow_{def} \forall \sigma\ \mathcal{M},\sigma \vDash \varphi \wedge Sent_{\mathscr{L}}(\varphi)$$

Existential quantifiers are implicit, and schematic letters such as $\mathcal{M},\sigma,\varphi,t$, and R_i^n are employed as variables, to indicate the kind of object they range over. We adopt this notation in what follows.

Since every model $\mathcal{M} = \langle d,\mathcal{I} \rangle$ of $\mathscr{L}$ is an interpretation of $\mathscr{L}$ in $\mathscr{L}_{\mathsf{ZF}}$, it induces the following translation $\tau_{\mathcal{M}} : \mathscr{L} \rightarrow \mathscr{L}_{\mathsf{ZF}}$:

- $\tau_{\mathcal{M}}(c) := \mathcal{I}(c)$;
- $\tau_{\mathcal{M}}(v) := v$;
- $\tau_{\mathcal{M}}(f_i^n(t_1,\ldots,t_n)) := \mathcal{I}(f_i^n)(\tau_{\mathcal{M}}(t_1),\ldots,\tau_{\mathcal{M}}(t_n))$;
- $\tau_{\mathcal{M}}(R_i^n t_1,\ldots,t_n) := \langle \tau_{\mathcal{M}}(t_1),\ldots,\tau_{\mathcal{M}}(t_n) \rangle \in \mathcal{I}(R_i^n)$;
- $\tau_{\mathcal{M}}(\neg\varphi) := \neg\tau_{\mathcal{M}}(\varphi)$;
- $\tau_{\mathcal{M}}(\varphi \wedge \psi) := \tau_{\mathcal{M}}(\varphi) \wedge \tau_{\mathcal{M}}(\psi)$;

- $\tau_{\mathcal{M}}(\forall v\varphi) := \forall v(v \in d \to \tau_{\mathcal{M}}(\varphi))$.

Proposition 8. *The following are theorems of* ZF*:*

- $(\mathcal{M} \vDash \varphi) \leftrightarrow \tau_{\mathcal{M}}(\varphi)$*, for every sentence* $\varphi \in \mathscr{L}$
- $\forall x \forall \mathcal{M}(T_{\mathscr{L}}(x, \mathcal{M}) \to Sent_{\mathscr{L}}(x))$

We can define semantic consequence and logical truth in terms of models and our definition of *truth-in-a-model*, as follows:

$\varphi \in \mathscr{L}$ is a *semantic consequence* of $\Gamma \subseteq \mathscr{L}$ $(\Gamma \vDash \varphi)$ iff, for every model $\mathcal{M}$ and assignment σ in $\mathcal{M}$, if $\mathcal{M}, \sigma \vDash \psi$ for every $\psi \in \Gamma$, then $\mathcal{M}, \sigma \vDash \varphi$.

$\varphi \in \mathscr{L}$ is a *logical truth* $(\vDash \varphi)$ iff, for every model $\mathcal{M}$, $\mathcal{M} \vDash \varphi$.

Theorem 9 (Soundness and Completeness)**.** *Let* $\Gamma \subseteq \mathscr{L}$ *and* $\varphi \in L$*. Then,* $\Gamma \vdash \varphi$ *iff* $\Gamma \vDash \varphi$.