MA5219 LOGIC AND FOUNDATION OF MATHEMATICS I

Lecture 5: Arithmetization of syntax

Tin Lok Wong

27 August, 2018

In this lecture, we describe one way of coding syntactic objects (e.g., terms, formulas) as natural numbers under which the usual syntactic properties/operations (e.g., whether a variable appears free in a formula, substitution) translate to recursive properties/operations on the natural numbers. This process is called *arithmetization*.

To see why arithmetization is relevant for the Incompleteness Theorems, let us go into the proof. Gödel's original proof resembles the Liar Paradox, except that truth is replaced by entailment.

Liar Paradox. Consider the sentence

This sentence is not true.

By definition, it is true if and only if it is not true.

Outline of Gödel's proof of his First Incompleteness Theorem. Let T be a 'suitable' $\mathcal{L}_{\mathrm{A}}(\exp)$ theory such that $T \models \mathrm{R}(\exp)$ and $\mathbb{N} \models T$. We want an $\mathcal{L}_{\mathrm{A}}(\exp)$ sentence σ satisfying

$$T \not\models \sigma \quad \text{and} \quad T \not\models \neg\sigma.$$

The plan is to produce an $\mathcal{L}_{\mathrm{A}}(\exp)$ sentence σ such that

$$\mathrm{R}(\exp) \quad \models \quad \sigma \leftrightarrow \text{“}T \not\models \sigma\text{”}. \tag{$*$}$$

Thus σ asserts 'this sentence is not semantically entailed by T' over $\mathrm{R}(\exp)$.

- Suppose $T \models \sigma$. Then

$$\begin{array}{lll} & T \models \text{“}T \not\models \sigma\text{”} & \text{by } (*)\text{, since } T \models \mathrm{R}(\exp); \\ \therefore & \mathbb{N} \models \text{“}T \not\models \sigma\text{”} & \text{since } \mathbb{N} \models T; \\ \therefore & T \not\models \sigma & \text{since } \mathbb{N} \text{ is the real world.} \end{array}$$

- Suppose $T \models \neg\sigma$. On the one hand, this implies $\mathbb{N} \models \neg\sigma$ as $\mathbb{N} \models T$, and so $\mathbb{N} \not\models \sigma$ by the truth definition. On the other hand, this implies

$$\begin{array}{lll} & T \models \text{“}T \models \sigma\text{”} & \text{by } (*)\text{, since } T \models \mathrm{R}(\exp); \\ \therefore & \mathbb{N} \models \text{“}T \models \sigma\text{”} & \text{since } \mathbb{N} \models T; \\ \therefore & T \models \sigma & \text{since } \mathbb{N} \text{ is the real world;} \\ \therefore & \mathbb{N} \models \sigma & \text{since } \mathbb{N} \models T. \end{array}$$

In both cases, we have a contradiction. Thus $T \not\models \sigma$ and $T \not\models \neg\sigma$, as required. Notice $\mathbb{N} \models \sigma$ by $(*)$, because the real world $\mathbb{N}$ is a model of $\mathrm{R}(\exp)$. □

What we need to execute this proof.

- Express “$T \models y$” as an $\mathcal{L}_{\mathrm{A}}(\exp)$ formula $\varphi(y)$.
- Construct an $\mathcal{L}_{\mathrm{A}}(\exp)$ sentence σ such that $\mathrm{R}(\exp) \models \sigma \leftrightarrow \neg\varphi(\underline{\sigma})$.

Symbol	(	)	v	$\top$	$=$	$\neg$	$\vee$	$\exists$	0	1	$+$	$\times$	exp	$<$
Digit	1	2	3	4	5	6	7	8	9	A	B	C	D	E

Table 5.1: Conversion between symbols in $\mathcal{L}_{\mathrm{A}}(\exp)$ and hexadecimal digits

The first bullet point will be dealt with in the next two lectures. In this lecture, we will deal with the second bullet point: arithmetization will allow us to form the numeral of any sentence, and the Diagonal Lemma will give us the required σ.

There are many ways to arithmetize syntax. Very often, exactly how one arithmetizes does not matter. So we choose one that is most convenient for us.

Every string of symbols, when stored on a computer, is a string of zeros and ones, and is thus coded as a number. In other words, we may read each string of symbols as a number written in a base bigger than the number of symbols in the language. There is however one problem: in the language there is an infinite supply of variable symbols $\mathsf{v}_0, \mathsf{v}_1, \mathsf{v}_2, \ldots$. The solution is to represent these as $\mathsf{v}, \mathsf{vv}, \mathsf{vvv}, \ldots$, so that only two symbols are needed. We employ hexadecimal (i.e., base 16) representation, where $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}$ stand for $0, 1, 2, \ldots, 15$ respectively.

Definition. Via Table 5.1, we view each string of symbols s in $\mathcal{L}_{\mathrm{A}}(\exp)$ as a natural number $\ulcorner s \urcorner$ written in hexadecimal representation in a one-to-one manner. We call $\ulcorner s \urcorner$ the *Gödel number* of s.

Example 5.1.
$$\begin{aligned} \ulcorner 2^{\mathsf{v}_0+1}\mathsf{v}_1 + 2^{\mathsf{v}_0} + \mathsf{v}_2 \urcorner &= \ulcorner (((\exp(\mathsf{v}+1) \times \mathsf{vv}) + \exp \mathsf{v}) + \mathsf{vvv}) \urcorner \\ &= 111\mathrm{D}13\mathrm{BA}2\mathrm{C}332\mathrm{BD}32\mathrm{B}3332 \quad \text{in hexadecimal} \\ &= 1 \times 16^{21} + 1 \times 16^{20} + \cdots + 3 \times 16^2 + 3 \times 16^1 + 2 \times 16^0. \end{aligned}$$

Via Gödel numbering, all the syntactic properties and operations we have met so far translate to recursive subsets and functions on $\mathbb{N}$. This can be seen using the Church–Turing Thesis.

Let us state and prove some closure properties of Σ_1 and Π_1 for future reference.

Definition. A Π_1 *formula* is an $\mathcal{L}_{\mathrm{A}}(\exp)$ formula of the form $\forall \bar{y}\ \theta$, where $\theta \in \Delta_0(\exp)$ and $\bar{y}$ is a possibly empty tuple of variables.

Definition. Two $\mathcal{L}_{\mathrm{A}}(\exp)$ formulas $\varphi(\bar{x}), \psi(\bar{x})$ are *semantically equivalent over an $\mathcal{L}_{\mathrm{A}}(\exp)$ theory T* if
$$T \models \forall \bar{x}\ \big(\varphi(\bar{x}) \leftrightarrow \psi(\bar{x})\big).$$
They are *semantically equivalent* if they are semantically equivalent over the empty theory $\varnothing$.

Lemma 5.2. Let $\theta, \eta \in \Sigma_1$ and $\theta', \eta' \in \Pi_1$. Then

(1) each of $\theta \wedge \eta, \theta \vee \eta, \neg\theta', \theta' \to \eta$ is semantically equivalent to a Σ_1 formula; and

(2) each of $\theta' \wedge \eta', \theta' \vee \eta', \neg\theta, \theta \to \eta'$ is semantically equivalent to a Π_1 formula.

Proof. The two parts are dual to each other. So we only prove (1). Let
$$\theta = \exists \bar{y}\ \theta_0, \quad \eta = \exists \bar{z}\ \eta_0, \quad \theta' = \forall \bar{y}'\ \theta'_0, \quad \eta' = \forall \bar{z}'\ \eta'_0,$$
where $\theta_0, \eta_0, \theta'_0, \eta'_0 \in \Delta_0(\exp)$. Applying some changes of variables if necessary, we may assume the tuples $\bar{y}, \bar{z}, \bar{y}', \bar{z}'$ are mutually disjoint. Then $\theta \wedge \eta, \theta \vee \eta, \neg\theta', \theta' \to \eta$ are respectively semantically equivalent to
$$\exists \bar{y}, \bar{z}\ (\theta_0 \wedge \eta_0), \quad \exists \bar{y}, \bar{z}\ (\theta_0 \vee \eta_0), \quad \exists \bar{y}'\ \neg\theta'_0, \quad \exists \bar{y}', \bar{z}\ (\theta'_0 \to \eta_0). \qquad \square$$

We introduce the following shorthand to make our notation slightly cleaner.

Definition. The *numeral* for a string of symbols s in $\mathcal{L}_{\mathrm{A}}(\exp)$, which we denote by $\underline{s}$, is the numeral for the Gödel number of s, i.e., we define $\underline{s} = \underline{\ulcorner s \urcorner}$.

The Diagonal Lemma is what allows a sentence to refer to itself as in the Liar Paradox. The proof is short, intriguing and surprisingly constructive.

Diagonal Lemma (also known as the Fixed-Point Theorem; Gödel)**.** For every $\mathcal{L}_{\mathrm{A}}(\exp)$ formula $\varphi(y)$, there is an $\mathcal{L}_{\mathrm{A}}(\exp)$ sentence σ such that

$$\mathrm{R}(\exp) \models \sigma \leftrightarrow \varphi(\underline{\sigma}).$$

Moreover, if $\varphi \in \Pi_1$, then we can require $\sigma \in \Pi_1$.

Proof. Let us first describe informally how the argument goes, without worrying too much about the details like the corners $\ulcorner \cdots \urcorner$ and the underlines $\underline{\cdots}$. Define the *diagonalization* of a formula $\theta(x)$ to be $\mathrm{Diag}(\theta) = \theta(\underline{\theta})$, i.e., the result obtained from $\theta(x)$ by replacing each free occurrence of the variable x with the term $\underline{\theta}$; see the next lecture for a precise definition. Consider the formula $\theta(x) = \varphi(\mathrm{Diag}(x))$. Then

$$\begin{aligned} \mathrm{Diag}(\theta) &= \theta(\underline{\theta}) && \text{by the definition of Diag;} \\ &= \varphi(\mathrm{Diag}(\theta)) && \text{by the definition of } \theta, \end{aligned}$$

so that we can set $\sigma = \mathrm{Diag}(\theta)$.

Now let us proceed formally. We will define a function $D\colon \mathbb{N} \to \mathbb{N}$ such that if $\theta(x)$ is an $\mathcal{L}_{\mathrm{A}}(\exp)$ formula, then $D(\ulcorner\theta\urcorner) = \ulcorner\theta(\underline{\theta})\urcorner$. More specifically, for each $m \in \mathbb{N}$, define

$$D(m) = \begin{cases} \ulcorner\theta(\underline{\theta})\urcorner, & \text{if } m = \ulcorner\theta\urcorner \text{ for some } \mathcal{L}_{\mathrm{A}}(\exp) \text{ formula } \theta(x); \\ 0, & \text{otherwise.} \end{cases}$$

By the Church–Turing Thesis, the function D is recursive. So Theorem 4.4 gives a Σ_1 formula $\rho(x, y)$ which represents D over $\mathrm{R}(\exp)$, i.e., for all $m \in \mathbb{N}$,

$$\mathrm{R}(\exp) \models \rho(\underline{m}, \underline{D(m)}) \wedge \forall y\ \big(\rho(\underline{m}, y) \to y = \underline{D(m)}\big).$$

Let $\theta(x)$ be the $\mathcal{L}_{\mathrm{A}}(\exp)$ formula $\forall y\ \big(\rho(x, y) \to \varphi(y)\big)$. In the case when $\theta \in \Pi_1$, we can choose a semantically equivalent Π_1 formula instead, by Lemma 5.2(2). We verify that the sentence $\sigma = \theta(\underline{\theta})$ has the properties we want.

Take $M \models \mathrm{R}(\exp)$. Then

$$\begin{aligned} & M \models \sigma && \\ \Leftrightarrow\quad & M \models \theta(\underline{\theta}) && \text{by the definition of } \sigma; \\ \Leftrightarrow\quad & M \models \forall y\ \big(\rho(\underline{\theta}, y) \to \varphi(y)\big) && \text{by the definition of } \theta; \\ \Leftrightarrow\quad & M \models \varphi(\underline{D(\ulcorner\theta\urcorner)}) && \text{by the choice of } \rho, \text{ as } \underline{\theta} = \underline{\ulcorner\theta\urcorner}; \\ \Leftrightarrow\quad & M \models \varphi(\underline{\theta(\underline{\theta})}) && \text{by the definition of } D; \\ \Leftrightarrow\quad & M \models \varphi(\underline{\sigma}) && \text{by the definition of } \sigma. \end{aligned}$$

□

The Diagonal Lemma easily implies that there is no formula which picks out exactly the Gödel numbers of true sentences in any model of $\mathrm{R}(\exp)$. The proof resembles the Liar Paradox.

Undefinability of Truth (Tarski)**.** There are no $\mathcal{L}_{\mathrm{A}}(\exp)$ formula $\psi(y)$ and $M \models \mathrm{R}(\exp)$ such that for all $\mathcal{L}_{\mathrm{A}}(\exp)$ sentences σ,

$$M \models \sigma \leftrightarrow \psi(\ulcorner\sigma\urcorner).$$

Proof. Apply the Diagonal Lemma to $\varphi = \neg\psi$ to find an $\mathcal{L}_{\mathrm{A}}(\exp)$ sentence σ such that

$$\mathrm{R}(\exp) \models \sigma \leftrightarrow \neg\psi(\underline{\sigma}).$$

As $M \models \mathrm{R}(\exp)$, we know $M \models \sigma \leftrightarrow \neg\psi(\underline{\sigma})$. By the truth definition, either $M \models \sigma$ or $M \models \neg\sigma$.

- If $M \models \sigma$, then $M \models \neg\psi(\ulcorner\sigma\urcorner)$.
- If $M \models \neg\sigma$, then $M \models \psi(\ulcorner\sigma\urcorner)$.

Either way, we have $M \not\models \sigma \leftrightarrow \psi(\ulcorner\sigma\urcorner)$. □

Assignment 5.3. Modify the proof of Tarski's theorem on the undefinability of truth to show that there are no Σ_1 formula $\psi(y)$ and $M \models \mathrm{R}(\exp)$ such that for all Π_1 sentences σ,

$$M \models \sigma \leftrightarrow \psi(\ulcorner\sigma\urcorner). \qquad \text{[5 points]}$$

In the next two lectures, we will study the semantic entailment relation in more detail.