Gödeli meeldetuletus

Tarmo Uustalu

Teooriapäevad Arulas, 3.–5.2.2003
• **What is this about?** (Rich) languages with a decided intended interpretation, (powerful) theories in such languages, axiomatized (powerful) theories in such languages.

• **Definition:** A *language* $L$ is a first-order logical language with equality and a denumerable amount of non-logical individual, function and predicate symbols. We assume a fixed intended interpretation. This singles out a subset of all $L$-sentences, the set of *true* sentences.

$\models A$ means $A$ is true in the intended interpretation.

An *$L$-theory* $T$ is a subset of all $L$-sentences, these sentences are called $T$-theorems.

$\vdash_T A$ means $A$ is a $T$-theorem.

An *axiomatized $L$-theory* is a $L$-theory generated by a p.r. subset of all $L$-sentences (called *axioms*) and the inference rules of first-order logic.
• **Definition:** Let $T$ be a theory in a language $L$ (with fixed intended interpretation).

- $T$ is said to be **consistent** (kooskõlaline), if $\vdash_T A$ implies $\not\vdash_T \neg A$ (there are no more theorems than syntactically ok).

- $T$ is said to be **sound** (korrektne), if $\vdash_T A$ implies $\models_T A$ (there are no more theorems than semantically ok).

- $T$ is said to be **syntactically complete** (süntaktiliselt täielik), if $\not\vdash_T A$ implies $\vdash_T \neg A$ (there are no less theorems than syntactically ok).

- $T$ is said to be **semantically complete** (semantiliselt täielik), if $\not\vdash_T A$ implies $\not\models_T A$ (there are no less theorems than semantically ok).
• **Observation:** The semantic properties are stronger than the syntactic ones:
  – soundness implies consistency,
  – and semantic completeness implies syntactic completeness.

• **Observation:** The converses don’t hold in general, but:
  – consistency implies soundness under the assumption of semantic completeness,
  – and syntactic completeness implies semantic completeness under the assumption of soundness.

• $T$ syntactically perfect, if it’s both consistent and syntactically complete, then for every sentence $A$, either $\vdash_T A$ or $\vdash_T \neg A$ (which mimicks bivalence).

• $T$ is semantically perfect, if it’s both sound and semantically complete, then theoremhood exactly captures truth.
• **Definition:** A language $L$ is *rich* if natural numbers, p.r. operations on natural numbers and p.r. relations on natural numbers are effectively *represented* (faithfully wrt. the intended interpretation) in $L$ by terms, schematics and schematics sentences.

Terms representing natural numbers are called *numerals*.

• **Definition:** An $L$-theory $T$ is *powerful*, if natural numbers, p.r. operations and p.r. relations on them satisfy the following *presentation conditions* (see below):

  – for $f$ a p.r. operation,

    $$\vdash_T \bar{f}[\bar{m}_1, \ldots, \bar{m}_n] \equiv \bar{m} \text{ iff } f(m_1, \ldots, m_n) = m$$

  – for $p$ a p.r. relation,

    $$\vdash_T \bar{p}[\bar{m}_1, \ldots, \bar{m}_n] \text{ iff } p(m_1, \ldots, m_n)$$

($\bar{m}$ denotes the representation of $m$.)
• **Fact:** The terms and sentences (and schematic terms and schematic sentences) of any rich language $L$ (with denumerable signature) are effectively enumerable by natural numbers so that all important syntactic operations on them reduce to primitive recursive operations on numbers (*Gödel numbers*).

• **Consequence:** Because of the representability of natural numbers in $L$, both the terms and sentences of $L$ therefore translate to $L$-numerals (*codes*). $⌜m⌝$ denotes the code of $m$.

In powerful $L$-theories, facts about important operations and relations concerning codes are reflected quite well since the presentation conditions hold.

• **Convention:** From now on, saying “language”, we always mean a rich language, and saying “theory”, we always mean a powerful theory.
• **Diagonalization Lemma:** Given a language $L$, one can for any schematic $L$-sentence $P$ effectively find a sentence $S$ s.t. $\models S \equiv P[⌞S⌝]$ and, for any $L$-theory $T$, $\vdash_T S \equiv P[⌞S⌝]$.

• **Proof:** Instantiating schematic $L$-sentences with $L$-numerals is a p.r. operation, reduced to Gödel numbers thus a p.r. operation on numbers, hence representable in $L$.

Let $\text{subst}$ be the schematic $L$-term representing it. Then $\models \text{subst}[⌞Q⌝,t] = ⌞Q[t]⌝$ for any schematic $L$-sentence $Q$ and any numeral $t$. For an $L$-theory $T$, $\vdash_T \text{subst}[⌞Q⌝,t] \equiv ⌞Q[t]⌝$ by the presentation conditions.

Consider any schematic $L$-sentence $P$. Let $D$ be the diagonal schematic $L$-sentence given by $D[t] := P[\text{subst}[t,t]]$.

Set $S := D[⌞D⌝]$. Then

$$\models S \equiv P[⌞S⌝] \text{ and } \vdash_T S \equiv P[⌞S⌝]$$

since by the definitions of $S$ and $D$, $S \equiv P[⌞S⌝]$ is identical to $P[\text{subst}[⌞D⌝,⌞D⌝]] \equiv P[⌞D[⌞D⌝]⌝]$.
• **Tarski’s theorem about non-representability of truth.** Given a language $L$, truth of $L$-sentences is non-representable in $L$: there is no schematic $L$-sentence $\text{True}$ such that

$$\models A \iff \models \text{True}[\neg A]$$

• **Proof.** Suppose a schematic $L$-sentence $\text{True}$ with the stated property exists. Then, applying the Diagonalization Lemma to the schematic $L$-sentence $\neg \text{True}$, we can produce an $L$-sentence $\text{Tarski}$ such that $\models \text{Tarski} \equiv \neg \text{True}[\neg \text{Tarski}]$, the effect that $\models \text{Tarski} \iff \neg \models \text{True}[\neg \text{Tarski}]$, which, by our assumption about $\text{True}$, happens iff $\not\models \text{Tarski}$.

Hence $\text{Tarski}$ is a sentence stating its own falsity, a “liar”. Independent of whether $\text{Tarski}$ is true or false, it is true and false, which cannot be.
• **Gödel’s theorem about representability of theoremhood.** Given a language $L$, theoremhood in an *axiomatized* $L$-theory $T$ is effectively representable in $L$: one can effectively find a schematic sentence $\text{Thm}_T$ in $L$ s.t.

$$\vdash_T A \text{ iff } \models \text{Thm}_T[⌜A⌝]$$

• **Proof:** For an axiomatized $L$-theory $T$, the relation of a sequence of $L$-sentences being a $T$-proof of a $L$-sentence is a p.r. relation, reduced to Gödel numbers, thus a p.r. relation on numbers, thus effectively representable in $L$. Let $\text{Proof}_T$ be the schematic $L$-sentence representing it.

$\text{Thm}_T$ is constructed by letting $\text{Thm}_T[t] := \exists x. \text{Nat}[x] \land \text{Proof}_T[x,t]$. 

9
• **Lemma (Gödel):** Given a language $L$, each axiomatized $L$-theory $T$ satisfies the following derivability conditions (tuletatavustingimused):

**D1** $\vdash_T A$ implies $\vdash_T \text{Thm}_T[\neg A]$ (the theory is positively introspective),

**D2** $\vdash_T \text{Thm}_T[\neg A \supset B] \supset (\text{Thm}_T[\neg A] \supset \text{Thm}_T[\neg B])$ (the theory knows it's closed under modus ponens),

**D3** $\vdash_T \text{Thm}_T[\neg A] \supset \text{Thm}_T[\neg \text{Thm}_T[\neg A]]$ (the theory knows it is positively introspective).

• **Proof:** Hard work (unrewarding).
• **Corollary:** Given a language $L$, a sound **axiomatized** $L$-theory $T$ is necessarily semantically incomplete (and hence because of the assumption of soundness, also syntactically incomplete).

• **Proof:** If some $L$-theory $T$ was both sound and semantically complete, $T$-theoremhood of $L$-sentences would be the same as truth. But one is $L$-representable, the other is not.
• Gödel’s first incompleteness theorem: Given a language $L$, for an $L$-theory $T$, one can effectively find an $L$-sentence $\text{Godel}_T$ s.t.

- if $T$ is consistent, then $\not\vdash_T \text{Godel}_T$, but $\models \text{Godel}_T$ (so $T$ is semantically incomplete),

- if $T$ is omega-consistent, then $\not\vdash_T \neg\text{Godel}_T$ (so $T$ is also syntactically incomplete).

• Proof: For an axiomatized $L$-theory $T$, we know that a schematic $\text{Thm}_T$ exists s.t. $\vdash_T A$ iff $\models \text{Thm}_T[\neg A]$.

Using the Diagonalization Lemma, we construct $\text{Godel}_T$ as an $L$-sentence

$$\models \text{Godel}_T \equiv \neg \text{Thm}_T[\neg \text{Godel}_T] \text{ and } \vdash_T \text{Godel}_T \equiv \neg \text{Thm}_T[\neg \text{Godel}_T]$$

(so informally $\text{Godel}_T$ says it’s a non-$T$-theorem and that’s a $T$-theorem).

Assume $T$ is consistent. Suppose $\vdash_T \text{Godel}_T$. Then, by D1, also $\vdash_T \text{Thm}_T[\neg \text{Godel}_T]$. But then, by the construction of $\text{Godel}_T$, $\vdash_T \neg \text{Godel}_T$, which contradicts consistency.

Suppose $\not\models \text{Godel}_T$, then by the construction of $\text{Godel}_T$, $\models \text{Thm}_T[\neg \text{Godel}_T]$. But then, by the construction of $\text{Thm}_T$, equivalent to $\vdash_T \text{Godel}_T$, but we already have $\vdash_T \neg \text{Godel}_T$, so again we are contradicting consistency.
• **Remark:** Note that while Tarski is an antinomic sentence, it must not exist; Godel merely paradoxical, its existence looks potentially troublesome, but there is nothing harmful about it.
• Gödel’s second incompleteness theorem: Given a language \( L \), for an \( L \)-theory \( T \), if \( T \) is consistent, then

\[ \not \vdash_T \text{Cons}_T \]

where \( \text{Cons}_T := \neg \text{Thm}_T[\neg \bot] \) (which says \( T \) is consistent). (So consistent axiomatized theory \( T \) is not a \( T \)-theorem.)

• Proof:
Assume \( T \) is a consistent axiomatized \( L \)-theory. By the construction of \( \Gamma_{Godel} \), we have

\[ \vdash_T \text{Godel}_T \supset \neg \text{Thm}_T[\neg \text{Godel}_T] \]

From this, by D1, we get

\[ \vdash_T \text{Thm}_T[\neg \text{Godel}_T] \supset \neg \text{Thm}_T[\neg \text{Godel}_T] \]

from where, by D2, we further get

\[ \vdash_T \text{Thm}_T[\neg \text{Godel}_T] \supset \text{Thm}_T[\neg \text{Thm}_T[\neg \text{Godel}_T]] \]

But by D3 we also have

\[ \vdash_T \text{Thm}_T[\neg \text{Godel}_T] \supset \text{Thm}_T[\neg \text{Thm}_T[\neg \text{Godel}_T]] \]
Combining the last two using D2 and the construction of Cons$_T$, we get
\[ \vdash_T \text{Thm}_T[\neg \text{Godel}_T] \supset \neg \text{Cons}_T \]
which of course gives
\[ \vdash_T \text{Cons}_T \supset \neg \text{Thm}_T[\neg \text{Godel}_T] \]
Together with the construction of Godel$_T$ again (the second half of the equivalence), this yields
\[ \vdash_T \text{Cons}_T \supset \text{Godel}_T \]
If now it were the case that $\vdash_T \text{Cons}_T$, then also $\vdash_T \text{Godel}_T$, but since the First Incompleteness Theorem tell us the that $\not\vdash_T \text{Godel}_T$. 
15