Lecture 17

Gödel's Theorems

Completeness and Incompleteness Theorems

Javier Esparza
TUM

Source

G.S. Boolos, J.P. Burgess, R.C Jeffrey:
Computability and Logic. Cambridge University Press 2002.

Figure: Kurt Gödel (1906 - 1978)

A Hilbert Calculus for first-order logic

We take eight axiom schemes or axioms, with F, G as place-holders for formulas:

1. $F \rightarrow (G \rightarrow F)$
2. $(F \rightarrow (G \rightarrow H)) \rightarrow ((F \rightarrow G) \rightarrow (F \rightarrow H))$
3. $(\neg F \rightarrow \neg G) \rightarrow (G \rightarrow F)$
4. $F \rightarrow (\neg F \rightarrow G)$
5. $(\neg F \rightarrow F) \rightarrow F$
6. $\forall x\ F \rightarrow F[t/x]$ for every term t.
7. $\forall x\ (F \rightarrow G) \rightarrow (\forall x\ F \rightarrow \forall x\ G)$.
8. $F \rightarrow \forall x\ F$ if x does not occur free in F.

Derivations

An **instance** of an axiom is the result of substituting the place-holders of the axiom by formulas.

Easy to see: all instances are valid formulas.

Example: Instance of (4) with $\neg A \rightarrow B$ and $\neg C$ for F and G:

$$(\neg A \rightarrow B) \rightarrow (\neg(\neg A \rightarrow B) \rightarrow \neg C)$$

Let S be a (possibly empty) set of formulas - also called **hypothesis** - and let F be a formula.

We write $S \vdash F$ and say that F is a **syntactic consequence of** S if one of these conditions holds:

Axiom:	F is an instance of an axiom
Hypothesis:	$F \in S$
Modus Ponens:	$S \vdash G \rightarrow F$ and $S \vdash G$, i.e. both $G \rightarrow F$ and G are syntactic consequences of S.

Gödel's Completeness Theorem

Theorem (Gödel's Completeness Theorem)

Let S be a (possibly empty) set of hypothesis, and let F be a formula.

F is a syntactic consequence of S iff it is a consequence of S.

Corollary

A formula is valid iff it is a syntactic consequence of the empty set of formulas.

A corollary of the theorem

Lemma

If a theory T is axiomatizable and complete, then T is decidable.

Proof.

If T inconsistent then T contains all closed formulas, and the algorithm that answers "$F \in T$" for every input F decides T.
If T consistent, let A be the set of axioms of T (not the axioms of the Hilbert Calculus). We consider the following algorithm:

- Input: F
  Enumerate all syntactic consequences of A, and for each new syntactic consequence G do:
  - If $G = F$, then halt with "$F \in T$"
  - If $G = \neg F$, then halt with "$F \notin T$"

Observe: the syntactic consequences of S can be enumerated. We prove this algorithm is correct. □

Proof.

(Continued.)

- If the algorithm answers "$F \in T$", then $F \in T$.
  If the algorithm answers "$F \in T$", then F is a syntactic consequence of S, and so a consequence of A. Since T is a theory, $F \in T$.
- If algorithm answers "$F \notin T$", then $F \notin T$.
  If the algorithm answers "$F \in T$", then $\neg F$ is a consequence of A and so $\neg F \in T$. By consistency, $F \notin T$.
- The algorithm terminates.
  Since T is complete, either $F \in T$ or $\neg F \in T$.
  Assume w.l.o.g. $F \in T$.
  Since T is axiomatizable, F is a consequence of A.
  So F is a syntactic consequence of A.
  So eventually $G := F$ and the algorithm terminates.

□

The signature of arithmetic

The signature σ_A of arithmetic contains:

- a constant 0,
- a unary function symbol s,
- two binary function symbols $+$ and $\cdot$, and
- a binary predicate symbol $<$.

ARITH is the theory containing the set of closed formulas over σ_A that are true in the canonical structure.

ARITH contains "all the theorems of calculus".

More consequences

Theorem

ARITH *is undecidable.*

Proof.

By reduction from the halting problem, similar to undecidability proof for validity of predicate logic. □

Theorem

ARITH *is not axiomatizable.*

Proof.

Since ARITH is undecidable, consistent, and complete, it is not axiomatizable (see Lemma). □

Gödel's first incompleteness theorem

Theorem (Gödel's first incompleteness theorem)

Let $\mathcal{X}$ be any decidable set of formulas such that $\mathcal{X} \subseteq$ ARITH. Then the theory $T_{\mathcal{X}}$ (the theory of all formulas entailed by $\mathcal{X}$) is incomplete.

Proof.

Since ARITH is not axiomatizable, there is a formula $F \in$ ARITH such that $\mathcal{X} \not\models F$ and so $F \notin T_{\mathcal{X}}$.

Assume now $\neg F \in T_{\mathcal{X}}$. Then $\mathcal{X} \models \neg F$ and since $\mathcal{X} \subseteq$ ARITH we get $\neg F \in$ ARITH, contradicting $F \in$ ARITH.

So $F \notin T_{\mathcal{X}}$ and $\neg F \notin T_{\mathcal{X}}$, which proves that $T_{\mathcal{X}}$ is incomplete. □

Gödel's first incompleteness theorem

Observe: for every set $\mathcal{X}$ of axioms there is a fomula $F \in \textsc{Arith}$, i.e., F is true in the canonical structure, but F is not a syntactic consequence of $\mathcal{X}$ (unless some axiom of $\mathcal{X}$ is itself not true!)

In other words: for every set of true axioms, there are true formulas that cannot be deduced from the axioms.

But we have no idea how such formulas look like ...

Goal: given a set of axioms $\mathcal{X} \subseteq \textsc{Arith}$, construct a formula $F \in \textsc{Arith}$ such that $F \notin T_{\mathcal{X}}$

Minimal arithmetic

Minimal arithmetic Q is the axiom-based theory over σ_A having the following axioms:

$$
\begin{array}{lrl}
(Q1) & \forall x & \neg(0 = s(x)) \\
(Q2) & \forall x \forall y & s(x) = s(y) \to x = y \\
(Q3) & \forall x & x + 0 = x \\
(Q4) & \forall x \forall y & x + s(y) = s(x + y) \\
(Q5) & \forall x & x \cdot 0 = 0 \\
(Q6) & \forall x \forall y & x \cdot s(y) = (x \cdot y) + x \\
(Q7) & \forall x & \neg(x < 0) \\
(Q8) & \forall x \forall y & x < s(y) \leftrightarrow (x < y \vee x = y) \\
(Q9) & \forall x \forall y & x < y \vee x = y \vee y < x
\end{array}
$$

Peano arithmetic

Peano arithmetic P is the axiom-based theory over σ_A having Q1-Q9 as axioms plus all closed formulas of the form

$$\text{(I)} \quad \forall \mathbf{y}\, F(0,\mathbf{y}) \;\wedge\; \forall x\, (\, F(x,\mathbf{y}) \rightarrow F(s(x),\mathbf{y})\,) \rightarrow \forall x\, F(x,\mathbf{y})$$

where $\mathbf{y} = (y_1, \ldots y_n)$.

Observe: I is an axiom scheme; the set of axioms of P is infinite but decidable.

Some theorems of Q (and P)

$\neg(0 = s^n(0))$ for every $n \geq 1$

$\neg(s^n(0) = s^m(0))$ for every $n, m \geq 1$, $n \neq m$

$\forall x \;\; x < 1 \leftrightarrow x = 0$

$\forall x \;\; x < s^{n+1}(0) \leftrightarrow (x = 0 \vee x = s(0) \vee \ldots \vee x = s^n(0))$

$s^n(0) + s^m(0) = s^l(0)$ for every $n, m, l \geq 1$ such that $n + m = l$

$s^n(0) \cdot s^m(0) = s^l(0)$ for every $n, m, l \geq 1$ such that $n \cdot m = l$

Gödel encodings

A **Gödel encoding** is an injective function that maps every formula over σ_A to a natural number called its **Gödel number**.

Simple Gödel encoding:

- assign to each symbol of the formula its ASCII code, and
- assign to a formula the concatenation of the ASCII codes of its symbols.

Gödel encodings

Example (Wikipedia): the formula

$$x = y \to y = x$$

written in ASCII as

```
x=y => y=x
```

corresponds to the sequence

120-061-121-032-061-062-032-121-061-120

of ASCII codes, and so it is assigned the number

120061121032061062032121061120

Gödel's Gödel encoding

Let p_n denote the n-th prime number.

Gödel's encoding assigns to each symbol λ a number $g(\lambda)$, and to a sequence $\lambda_1 \cdots \lambda_n$ of symbols the number

$$2^{g(\lambda_1)} \cdot 3^{g(\lambda_2)} \cdot 5^{g(\lambda_3)} \cdot \ldots \cdot p_n^{g(\lambda_n)}$$

What are Gödel encodings good for?

A formula $F(x)$ over σ_A with a free variable x defines a **property of numbers**: the property satisfied exactly by the numbers n such that $F(s^n(0))$ is true in the canonical structure.

We can easily construct formulas $Even(x)$, $Prime(x)$, $Power\_of\_two(x)$ ...

Via the encoding formulas "are" numbers, and so a formula also defines a property of formulas!

$$\begin{array}{rcccc} \text{numbers} & \to & \text{formulas} & & \\ \text{formula } F(x) & \to & \text{set of numbers} & \to & \text{set of formulas} \end{array}$$

Going further ...

We can (less easily) construct formulas like

- $\mathit{First\_symbol\_is\_}\forall(x)$
- $\mathit{At\_least\_ten\_symbols}(x)$
- $\mathit{Closed}(x)$
- ...

that are true in the canonical structure for $x \mapsto s^n(0)$ iff the number n encodes a formula and the formula satisfies the corresponding property.

And even further ...

We can construct (even less easily) a formula

- $In\_Q(x)$

that is true in the canonical structure with $x \mapsto s^n(0)$ iff the number n encodes a closed formula F such that $F \in \mathrm{Q}$.

The reason is

$$F \in \mathrm{Q} \text{ iff } \mathrm{Q1}, \ldots, \mathrm{Q9} \models F \text{ iff } \mathrm{Q1}, \ldots, \mathrm{Q9} \vdash F$$

and the derivation procedure amounts to symbol manipulation.

Same for any other set $\mathcal{X}$ of axioms.

Diagonal Lemma

Recall our goal: Given a set of axioms $\mathcal{X} \subseteq \text{ARITH}$, construct a formula $F \in \text{ARITH}$ such that $F \notin T_{\mathcal{X}}$

Let $\underline{F}$ denote the term $s^n(0)$ where n is the Gödel encoding of the formula F.

Intuition: $\underline{F}$ is a “name” we give to F

Lemma (Diagonal Lemma)

Let $\mathcal{X}$ be any set of axioms containing $Q1, \ldots Q9$. For every formula $B(y)$ there is a closed formula G such that $G \leftrightarrow B(\underline{G}) \in T_{\mathcal{X}}$.

We call G the **Gödel formula** of $B(x)$.

We have: G true i.t.c.s if and only if G has property B

Intuition: G asserts that G has property B (true or false in the canonical structure!)

Reaching the goal

Theorem

Let $\mathcal{X}$ be any set of axioms containing Q1, ... Q9.
Let $G_{\mathcal{X}}$ be the Gödel formula of $\neg In\_T_{\mathcal{X}}(x)$. Then $G_{\mathcal{X}} \in \textsc{Arith} \setminus T_{\mathcal{X}}$.

Proof.

(Idea.) By definition, $G_{\mathcal{X}}$ is true i.t.c.s iff $G_{\mathcal{X}} \notin T_{\mathcal{X}}$.

If $G_{\mathcal{X}}$ is false i.t.c.s. then $G_{\mathcal{X}} \in T_{\mathcal{X}}$.

Since $\mathcal{X} \subseteq \textsc{Arith}$, we have $G_{\mathcal{X}} \in \textsc{Arith}$.

But then, by definition of $\textsc{Arith}$, $G_{\mathcal{X}}$ is true i.t.c.s.

Contradiction!

So $G_{\mathcal{X}}$ is true i.t.c.s., i.e., $G_{\mathcal{X}} \in \textsc{Arith}$.

But then $G_{\mathcal{X}} \notin T_{\mathcal{X}}$, and so $G_{\mathcal{X}} \in \textsc{Arith} \setminus T_{\mathcal{X}}$. Done! □

Gödel's second incompleteness theorem

For any set of axioms $\mathcal{X}$ containing Q1 we have $0 = s(0) \notin T_{\mathcal{X}}$, and so $T_{\mathcal{X}}$ is consistent iff $0 = s(0) \notin T_{\mathcal{X}}$.

The **consistency formula** for $\mathcal{X}$ is the formula $\neg In\_T_{\mathcal{X}}(\underline{0=s(0)})$

Intuition: The consistency formula for $\mathcal{X}$ states that $T_{\mathcal{X}}$ is consistent.

Theorem (Gödel's second incompleteness theorem)

Let $\mathcal{X}$ be any set of axioms containing P. Then the consistency formula for $\mathcal{X}$ does not belong to $T_{\mathcal{X}}$.

Intuition: the consistency of a theory cannot be derived from the axioms of the theory.

Proving the Diagonal Lemma: Diagonalization

Let $F(x)$ be a formula with a free variable x.

The **diagonalization** of F is the closed formula

$$DiagF := \exists x\ x = \underline{\mathrm{F}} \wedge F(x)$$

Intuition: *DiagF* asserts that F has property F

Observe: *DiagF* and $F(\underline{\mathrm{F}})$ are logically equivalent, but they have different Gödel numbers.

The representation theorem

Theorem

There is a formula Diag(x, y) such that for every formula F

$$\forall y \;\; Diag(\underline{F}, y) \leftrightarrow y = \underline{DiagF}$$

can be derived in Q (and so in P).

Observe: the theorem does not hold for every set of axioms. For instance, it does not hold for the system Q1-Q4, since in that system we cannot infer anything about the product function.

Proof of the Diagonal Lemma

Lemma

Let $\mathcal{X}$ be any set of axioms containing Q1, ... Q9.
For every formula $B(y)$ there is a closed formula G such that $G \leftrightarrow B(\underline{G}) \in T_{\mathcal{X}}$.

Proof.

Let $A(x) := \exists y\ (Diag(x, y) \wedge B(y))$ and let $G := DiagA$.

Intuition: G asserts that the diagonalization of A (the formula asserting that A satisfies A) satisfies B.

Explicitely:

$$G := \exists x\ (x = \underline{A} \wedge A(x)) := \exists x\ (x = \underline{A} \wedge \exists y\ (Diag(x, y) \wedge B(y)))$$

(To be continued) □

Proof of the Diagonal Lemma

Continued.

The formula $G \leftrightarrow \exists y\ (Diag(\underline{\mathrm{A}}, y) \wedge B(y))$ is valid, and so, since valid formulas belong to every theory, we have

$$G \leftrightarrow \exists y\ (Diag(\underline{\mathrm{A}}, y) \wedge B(y)) \ \in \ T_{\mathcal{X}}$$

Since $G := DiagA$, we have by the representation theorem:

$$\forall y\ (Diag(\underline{\mathrm{A}}, y) \leftrightarrow y = \underline{\mathrm{G}}) \ \in \ T_{\mathcal{X}}$$

And so, since $T_{\mathcal{X}}$ is closed under consequence, we get

$$G \leftrightarrow \exists y\ (y = \underline{\mathrm{G}} \ \wedge \ B(y)) \ \in \ T_{\mathcal{X}}$$

and for the same reason

$$G \leftrightarrow B(\underline{\mathrm{G}}) \ \in \ T_{\mathcal{X}}$$

□