LSS 2018: Computability and Incompleteness

5. Provability Predicates
Gödel's Second Incompleteness Theorem

Michael Norrish

Michael.Norrish@data61.csiro.au

Outline

Last Time. . .

Representability

- All recursive functions are representable in extensions of $\mathcal{Q}$

Arithmetic Cannot be Captured

- The diagonalisation function is computable
- So any candidate "theorem-hood" notion can be turned against itself
  - "I am true iff I am not a theorem"
- Truth is not definable in arithmetic (Tarski)
- Arithmetic is not axiomatisable (Gödel)

Peano Arithmetic

Called, variously: PA (*Johnstone*), $\mathcal{Z}$ (*B&J*), $\mathcal{S}$ (*Mendelson*).

Take $\mathcal{Q}$, and add induction:

- If P is a formula with x free, then the universal closure of

  $$P(0) \land (\forall m.\ P(m) \Rightarrow P(s(m))) \Rightarrow (\forall n.\ P(n))$$

  is an axiom.

  (Where $P(a)$ means P with x replaced by a.)

The result is a formal system with an infinite number of axioms.

- However, the axioms are still decidable.

Outline

Proofs are Computably Checkable

A proof in a formal system is a sequence of formulas such that every formula in the sequence is

- an instance of an axiom; or
- is the result of applying a rule of inference to one or more formulas earlier in the sequence

For human consumption, we usually indicate a non-axiom's forebears explicitly.

But we could just check all possible earlier formulas.

Proofs are Arithmetisable

Already know how to map

- formulas into numbers
- lists of numbers into numbers.

Can therefore turn a proof into a number.

Checking this number is really a proof is computable, hence representable in extensions of $\mathcal{Q}$.

Given formula A, can also check that the last formula in a proof is equal to A.

Thus

$$Proof(p, \ulcorner A \urcorner) = p \text{ is a proof of } A$$

is definable.

A Provability Predicate

Let $\mathit{Provable}(n) \stackrel{\text{def}}{=} (\exists p.\ \mathit{Proof}(p, n))$

Write $\Box A$ for $\mathit{Provable}(\ulcorner A \urcorner)$.

Important Properties of Provability:

- if $\vdash A$ then $\vdash \Box A$
- $\vdash \Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B)$
- $\vdash \Box A \Rightarrow \Box(\Box A)$

In $\mathcal{Z}$ the above can all be proved; as can

- if $\vdash_{\mathcal{Z}} \Box A$ then $\vdash_{\mathcal{Z}} A$

Provability Does Not Define Theorem-Hood

Last time, we proved the indefinability of theorem-hood.

Definability required

$$\begin{array}{lcll} \vdash_{\mathcal{T}} Thm(nt(n)) & \text{iff} & \vdash_{\mathcal{T}} gn^{-1}(n) & \quad (1) \\ \vdash_{\mathcal{T}} \neg Thm(nt(n)) & \text{iff} & \nvdash_{\mathcal{T}} gn^{-1}(n) & \quad (2) \end{array}$$

Provability ($\Box$) only gives us (1).

So what happens if we replay the proof of indefinability with $\Box$?

The Gödel Sentence

We have a G such that $\vdash_{\mathcal{Z}} G \iff \neg\Box G$ (1)

- This is the Gödel sentence for our theory.

We also know that $\vdash_{\mathcal{Z}} G$ iff $\vdash_{\mathcal{Z}} \Box G$ (2)

If $\mathcal{Z}$ is consistent, then:

G is not a theorem of $\mathcal{Z}$.

- If it were, then $\vdash_{\mathcal{Z}} G$. So, $\vdash_{\mathcal{Z}} \Box G$ by (2). But also, $\vdash_{\mathcal{Z}} \neg\Box G$ by (1), making Z inconsistent.

$\neg G$ is not a theorem of $\mathcal{Z}$.

- If it were, then $\vdash_{\mathcal{Z}} \Box G$ by (1). Then $\vdash_{\mathcal{Z}} G$ by (2).
  Again making $\mathcal{Z}$ inconsistent.

Gödel's First Incompleteness Theorem Concretely

As long as our logic $\mathcal{T}$ is strong enough to give us

$$\vdash_{\mathcal{T}} G \quad \text{iff} \quad \vdash_{\mathcal{T}} \Box G$$

we know

If $\mathcal{T}$ is consistent, then $\nvdash_{\mathcal{T}} G$ and $\nvdash_{\mathcal{T}} \neg G$

In other words, G demonstrates $\mathcal{T}$'s incompleteness.

Moreover, we do know that $\vdash_{\mathcal{T}} G \iff \neg\Box G$

- This says that G is true iff G is not provable.
- Having just proved G's unprovability, we can conclude G is true.

Henkin's Formula

On one hand, G says that G isn't derivable.

Diagonalisation also gives us H such that

$$\vdash_{\mathcal{T}} H \iff \Box H$$

or

H says that H is derivable

But is H true?

Löb's Theorem

By far the weirdest result of the course:

If $\vdash_{\mathcal{T}} \Box A \Rightarrow A$, *then* $\vdash_{\mathcal{T}} A$

Can also write:

$$\Box(\Box A \Rightarrow A) \Rightarrow \Box A$$

which is the the axiom for modal provability logic.

(Why does provability "correspond" to a binary relation that is transitive and well-founded?)

Proof of Löb's Theorem

Theorem: if $\vdash_{\mathcal{T}} \Box A \Rightarrow A$, then $\vdash_{\mathcal{T}} A$

Diagonalise formula $\Box x \Rightarrow A$, giving L such that

1	$\vdash_{\mathcal{T}} L \iff (\Box L \Rightarrow A)$	
2	$\vdash_{\mathcal{T}} L \Rightarrow (\Box L \Rightarrow A)$	(bicond elimination)
3	$\vdash_{\mathcal{T}} \Box(L \Rightarrow (\Box L \Rightarrow A))$	(PP1)
4	$\vdash_{\mathcal{T}} \Box L \Rightarrow \Box(\Box L \Rightarrow A)$	(PP2)
5	$\vdash_{\mathcal{T}} \Box L \Rightarrow (\Box\Box L \Rightarrow \Box A)$	(PP2 on right)
6	$\vdash_{\mathcal{T}} \Box L \Rightarrow \Box A$	(PP3 eliminates $\Box\Box L$)
7	$\vdash_{\mathcal{T}} \Box L \Rightarrow A$	($\Box A \Rightarrow A$ by assumption)
8	$\vdash_{\mathcal{T}} L$	(7,1)
9	$\vdash_{\mathcal{T}} \Box L$	(PP1)
10	$\vdash_{\mathcal{T}} A$	(7,9)

Löb's Theorem Proves the Henkin Sentence

Henkin sentence is $\vdash_{\mathcal{T}} H \iff \Box H$

If that's provable, so too is $\vdash_{\mathcal{T}} \Box H \Rightarrow H$.

By Löb's Theorem: $\vdash_{\mathcal{T}} H$

So the sentence that "says of itself that it is provable", is indeed true.

Outline

Provability Gives us Arithmetisation of Consistency

Write $\bot$ for $\mathbf{0} \neq \mathbf{0}$. (Recall that $\vdash \bot \Rightarrow A$ for any A.)

Write $\mathsf{Con}_{\mathcal{T}}$ for $\neg\Box\bot$ ("false" is not provable).

- Consistency was "actually" simultaneous derivation of A and $\neg A$ for some A
- But the two are equivalent.

Consistency is Unprovable (Sketchy Version)

Want to show

$$\vdash_{\mathcal{T}} \mathsf{Con}_{\mathcal{T}} \Rightarrow G$$

Then, $\mathsf{Con}_{\mathcal{T}}$ can't be derivable, because if it were, G would be too.

We know that G "means" 'G is not derivable'.

Gödel's First Incompleteness Theorem says

If $\mathcal{T}$ is consistent, then G is not derivable.

But that's just what we want to prove!

- Just have to be able to carry out proof of Gödel's First Incompleteness Theorem in $\mathcal{T}$ **Done**

Consistency is Unprovable (Löb Version)

Suppose we did have $\vdash_{\mathcal{T}} \mathsf{Con}_{\mathcal{T}}$, or $\vdash_{\mathcal{T}} \neg\Box\bot$.

Then get: $\vdash_{\mathcal{T}} \Box\bot \Rightarrow \bot$

- by propositional principle of proving anything from a false assumption

Löb's Theorem then says $\vdash_{\mathcal{T}} \bot$ (false is derivable after all!)

A contradiction, so consistency is not provable. **Done**

Consistency is Unprovable (non-Löb PP Version)

Recall that G demonstrates $\mathcal{T}$'s incompleteness (is unprovable).

Now want to argue that if $\mathcal{T}$ extends $\mathcal{Z}$, then

$$\vdash_{\mathcal{T}} \mathsf{Con}_{\mathcal{T}} \Rightarrow G$$

(if $\mathsf{Con}_{\mathcal{T}}$ were provable, G would be too).

- Have (provability property): $\vdash_{\mathcal{T}} \Box G \Rightarrow \Box\Box G$
- Thus (diagonal property of G): $\vdash_{\mathcal{T}} \Box G \Rightarrow \Box\neg G$
  - "if I can prove G, then I can also prove $\neg G$"
- So, $\vdash_{\mathcal{T}} \Box G \Rightarrow \Box\bot$
- Diagonal property of G: $\vdash_{\mathcal{T}} \neg G \Rightarrow \Box\bot$
- Contrapositively: $\vdash_{\mathcal{T}} \neg\Box\bot \Rightarrow G$

Done

Gödel's Second Incompleteness Theorem

If $\mathcal{T}$ is at least as powerful as $\mathcal{Z}$, then it cannot simultaneously:

- Be consistent
- Prove its own consistency

Outline

Would a Consistency Proof of $\mathcal{T}$ in $\mathcal{T}$ Be Convincing?

Imagine we are doubtful about $\mathcal{T}$.

A consistency proof would be reassuring.

But if that proof is carried out in $\mathcal{T}$ too,
how does that assuage our doubts?

- If it could be done in a small part of $\mathcal{T}$, maybe…

Consistency is Possible by Other Means

Peano Arithmetic was proved consistent by Gentzen.
($\mathcal{Q}$'s consistency follows too.)

He didn't do it in PA, but used a different logical system.

Nor was his system stronger than PA; just different.

Yikes, An Infinite Regress Awaits!

If we can't prove our interesting systems consistent except by recourse to other systems, this is a neverending process!

Yikes, An Infinite Regress Awaits!

If we can't prove our interesting systems consistent except by recourse to other systems, this is a neverending process!

So what?

- We have the same problem whenever we set up our logical systems; we have to start with some set of axioms.
- "We don't need Gödel to tell us that we cannot accept a proof in one formal system only on the basis of proof in another formal system."—Franzén

Note

Consistent systems don't have to prove true theorems.

My Own Self-Doubt-Casting Sentence

If anyone says

“X because of Gödel’s Theorem”

or

“Thanks to Gödel’s Theorem, X”

or variants of the same...

...they’re talking nonsense.

My Own Self-Doubt-Casting Sentence

If anyone says

“X because of Gödel’s Theorem”

or

“Thanks to Gödel’s Theorem, X”

or variants of the same…

…they’re talking nonsense. (To a first approximation.)

Examples from Franzén

- *Religious people claim that all answers are to be found in the Bible or in whatever text they use. That means the Bible is a complete system, so Gödel seems to indicate it cannot be true. And the same may be said of any religion which claims, as they all do, a final set of answers.*
- *As Gödel demonstrated, all consistent formal systems are incomplete, and all complete formal systems are inconsistent. The U.S. Constitution is a formal system, after a fashion. The Founders made the choice of incompleteness over inconsistency, and the Judicial Branch exists to close that gap of incompleteness.*
- *Gödel demonstrated that any axiomatic system must be either incomplete or inconsistent, and inasmuch as Ayn Rand's philosophy of Objectivism claims to be a system of axioms and propositions, one of these two conditions must apply.*
- *Nonstandard models and Gödel's incompleteness theorem point the way to God's freedom to change both the structure of knowing and the objects known.*

Mathematics Floundering in a Relativistic Sea?

We can extend $\mathcal{T}$ by adding either G or $\neg G$ as a new axiom.

The resulting theory will be consistent if $\mathcal{T}$ was.

- How do we pick which one to take?

For $\mathcal{Z}$ (PA), we know that $G \iff \mathsf{Con}_{\mathcal{Z}}$.

- We also know $\mathsf{Con}_{\mathcal{Z}}$ (Gentzen), so we should pick $\mathcal{Z} + \mathcal{G}$.

For more complicated systems (*e.g.*, ZFC set theory), "ordinary mathematics" does not necessarily know their consistency.

- but systems $\mathsf{ZFC} + \neg\mathsf{Con}_{\mathsf{ZFC}}$ are uninteresting

Gödel and AI

Lucas:

> *However complicated a machine we construct, it will, if it is a machine, correspond to a formal system, which in turn will be liable to the Gödel procedure for finding a formula unprovable in that system. This formula the machine will be unable to produce as true, although a mind can see that it is true.*

False.

- The Gödel formula is equivalent to the consistency of the system; it is not true in general.
- The "human mind" is not known to have any special ability to determine the consistency of arbitrary formal systems.

Also, see Franzén for more on Penrose's various arguments.

Summary

Provability Predicates

- Logical theories as strong as $\mathcal{Z}$ can capture the notion of provability.
- Modal axioms must characterise the putative modality ($\Box$)
- Löb: if $\vdash_{\mathcal{T}} \Box A \Rightarrow A$, then $\vdash_{\mathcal{T}} A$

Gödel's Second Incompleteness Theorem

- A system as strong as $\mathcal{Z}$ cannot both be consistent and prove its own consistency.

Be Careful Out There

Course Summary

Computability

- Turing Machines and Recursive Functions are equivalent.
  - No extant computational model is more powerful
- Uncomputable problems exist (Halting Problem, notably)

Logic and Incompleteness

- Validity in FOL is undecidable (by reduction to Halting Problem)
- Logics with minimal arithmetic can represent computable functions.
- By diagonalisation of formulas (a computable procedure):
  - arithmetic truth is undecidable;
  - no theory can be all three of consistent, complete, axiomatisable
- No theory extending $\mathcal{Z}$ can prove its own consistency