

# Provability Logic (from Gödel 1933 to the 21st century)

Amílcar Sernadas

Centro de Lógica e Computação - Instituto Superior Técnico

April 28, 2006

# Plan

- Gödel's seminal proposal
- Löb's answer to Henkin's question
- Gödel-Löb modal system of provability (*GL*)
- *GL* versus Peano arithmetic
- Recent developments

# Gödel's seminal proposal

## Gödel, 1933

"An interpretation of the intuitionistic propositional calculus" published in the proceedings of Karl Menger's mathematical colloquium at the University of Vienna for 1931-32.

Gödel, K., "Eine Interpretation des Intuitionistischen Aussagenkalküls," *Ergebnisse eines Mathematischen Kolloquiums*, Vol. 4 (1933): 39-40.

"An Interpretation of the Intuitionistic Propositional Calculus" in Gödel, K., *Collected Works*, ed. by Feferman, S. et al., Oxford and New York: Oxford University Press (1995), Vol. 3: 296-302.

## Provability connective

Let  $\mathbf{B} \alpha$  mean that  $\alpha$  is provable (*beweisbar*). To this end, the propositional calculus should be enriched with the axioms:

- $\mathbf{B} \alpha \Rightarrow \alpha$ ;
- $\mathbf{B} \alpha \Rightarrow (\mathbf{B}(\alpha \Rightarrow \beta) \Rightarrow \mathbf{B} \beta)$ ;
- $\mathbf{B} \alpha \Rightarrow \mathbf{B} \mathbf{B} \alpha$ ;

and the inference rule:

- from  $\alpha$  infer  $\mathbf{B} \alpha$ .

## Interpretation of intuitionistic logic

Intuitionistic logic (Heyting, 1930) can be interpreted in the **B**-calculus, for instance, as follows:

Intuitionistic logic	<b>B</b> -calculus
$p$	$p$
$\neg \alpha$	$\neg \mathbf{B} \alpha$
$\alpha \Rightarrow \beta$	$\mathbf{B} \alpha \Rightarrow \mathbf{B} \beta$
$\alpha \vee \beta$	$\mathbf{B} \alpha \vee \mathbf{B} \beta$
$\alpha \wedge \beta$	$\alpha \wedge \beta$

## Gödel's claims

- Translations of intuitionistic theorems derivable in the **B**-calculus;
- Presumably, the converse holds;
- The translation of  $p \vee \neg p$  is not derivable;
- A formula of the type  $\mathbf{B} \alpha \vee \mathbf{B} \beta$  is derivable only when one of the  $\mathbf{B} \alpha$  and  $\mathbf{B} \beta$  is derivable.

These claims were proved much later (McKinsey and Tarski, 1948).

## Goldblatt, 2005

*Those familiar with later developments will recognise the pregnancy of this brief note of scarcely more than a page.*

*Its translations provided an important connection between intuitionistic and modal logic that contributed to the development both of topological interpretations and of Kripke semantics for intuitionistic logic.*

*Its ideas also formed the precursor to the substantial branch of modal logic concerned with the modality “it is provable in Peano arithmetic that”.*

*It is now standard practice to present modal logics in the axiomatic style of Gödel (as extensions of ordinary propositional logic).*

# Löb's answer to Henkin's question

## Henkin's question, 1952

Inspired by the incompleteness theorems (Gödel, 1931), Henkin posed a deceptively simple question:

What can be said about sentences asserting their own provability?

More precisely, assuming that *PA* (Peano arithmetic) proves

$$\alpha \Leftrightarrow \text{Prov}([\alpha])$$

what can we say about  $\alpha$ ?

## Löb's surprising answer, 1955

Theorem: If  $\vdash_{PA} \text{Prov}(\ulcorner \alpha \urcorner) \Rightarrow \alpha$  then  $\vdash_{PA} \alpha$

That is, *PA* proves

$$\text{Prov}(\ulcorner \alpha \urcorner) \Rightarrow \alpha$$

only in the trivial case where *PA* already proves  $\alpha$  itself.

*PA* could not be more modest about its own veracity!

## Löb's conditions, 1955

- LC1:  $\vdash_{PA} \text{Prov}(\ulcorner \alpha \Rightarrow \beta \urcorner) \Rightarrow (\text{Prov}(\ulcorner \alpha \urcorner) \Rightarrow \text{Prov}(\ulcorner \beta \urcorner))$ ;
- LC2:  $\vdash_{PA} \text{Prov}(\ulcorner \alpha \urcorner) \Rightarrow \text{Prov}(\ulcorner \text{Prov}(\ulcorner \alpha \urcorner) \urcorner)$ ;
- LC3: If  $\vdash_{PA} \alpha$  then  $\vdash_{PA} \text{Prov}(\ulcorner \alpha \urcorner)$ .

(improving on Hilbert and Bernays, 1939)

Using Löb's conditions, Löb's theorem is derived from

Gödel, 1931

**Diagonal Lemma:** For every arithmetical formula  $\alpha$  with a single free variable  $x$  there is an arithmetical sentence  $\beta$  such that

$$\vdash_{PA} \beta \Leftrightarrow \alpha_{[\beta]}^x.$$

as follows:

- 1 By the Diagonal Lemma applied to  $Prov(x) \Rightarrow \alpha$ , there is a sentence  $\beta$  such that  $\vdash_{PA} \beta \Leftrightarrow (Prov(\ulcorner \beta \urcorner) \Rightarrow \alpha)$ .
- 2 Using LC3 and LC1 and some propositional reasoning we get  $\vdash_{PA} Prov(\ulcorner \beta \urcorner) \Rightarrow Prov(\ulcorner Prov(\ulcorner \beta \urcorner) \Rightarrow \alpha \urcorner)$ .
- 3 Using LC1 again, we get  $\vdash_{PA} Prov(\ulcorner \beta \urcorner) \Rightarrow (Prov(\ulcorner Prov(\ulcorner \beta \urcorner) \urcorner)) \Rightarrow Prov(\ulcorner \alpha \urcorner)$ .
- 4 By LC2 we know that  $\vdash_{PA} Prov(\ulcorner \beta \urcorner) \Rightarrow Prov(\ulcorner Prov(\ulcorner \beta \urcorner) \urcorner)$ .
- 5 Tautologically from 3 and 4,  $\vdash_{PA} Prov(\ulcorner \beta \urcorner) \Rightarrow Prov(\ulcorner \alpha \urcorner)$ .
- 6 Therefore, from  $\vdash_{PA} Prov(\ulcorner \alpha \urcorner) \Rightarrow \alpha$  we can conclude using 5 that  $\vdash_{PA} Prov(\ulcorner \beta \urcorner) \Rightarrow \alpha$ .
- 7 Hence, using 1 again, we get  $\vdash_{PA} \beta$  and, so, by LC3 we obtain  $\vdash_{PA} Prov(\ulcorner \beta \urcorner)$ .
- 8 Finally, from 6 and 7 by MP, we get  $\vdash_{PA} \alpha$ .

## Gödel's second incompleteness theorem as a corollary of Löb's theorem

If  $PA$  is consistent then  $\not\vdash_{PA} \neg Prov(\ulcorner \perp \urcorner)$ .

- 1 Assume  $\vdash_{PA} \neg Prov(\ulcorner \perp \urcorner)$ .
- 2 Then, tautologically,  $\vdash_{PA} Prov(\ulcorner \perp \urcorner) \Rightarrow \perp$ .
- 3 Using Löb's theorem, from 2 we get  $\vdash_{PA} \perp$ .
- 4 Hence,  $PA$  is not consistent.

## Modal logic of provability - System *GL*

Löb's characterization of provability seems to cry out for an investigation of:

### Provability as a modality

$\Box\alpha$  representing  $Prov(\ulcorner\alpha\urcorner)$

This effort was carried out much later, starting in the nineteen seventies with work by de Jongh, Kripke, Sambin, Segerberg, Smoryński, Solovay et al.

Fix a set  $P$  of propositional symbols.

## Language

The set  $L_{GL}$  of formulae is inductively defined as follows:

- $p \in L_{GL}$  for each  $p \in P$ ;
- $\neg \alpha \in L_{GL}$  if  $\alpha \in L_{GL}$ ;
- $\alpha \Rightarrow \beta \in L_{GL}$  if  $\alpha, \beta \in L_{GL}$ ;
- $\Box \alpha \in L_{GL}$  if  $\alpha \in L_{GL}$ .

Other propositional connectives introduced as abbreviations.  
Furthermore:

$\Diamond \alpha$  abbreviates  $\neg \Box \neg \alpha$

## Calculus

- Axioms:
  - Tautologies;
  - Normality:  $\Box(\alpha \Rightarrow \beta) \Rightarrow (\Box\alpha \Rightarrow \Box\beta)$ ;
  - Löb axiom:  $\Box(\Box\alpha \Rightarrow \alpha) \Rightarrow \Box\alpha$ ;
- Inference rules:
  - Modus Ponens: from  $\alpha$  and  $\alpha \Rightarrow \beta$  infer  $\beta$ ;
  - Necessitation: from  $\alpha$  infer  $\Box\alpha$ .

## Derivation

$$\Gamma \vdash_{GL} \alpha$$

if it is possible in a finite number of steps to infer  $\alpha$  from the axioms and the elements of  $\Gamma$  using the inference rules.

## Theoremhood

$$\vdash_{GL} \alpha \text{ if } \emptyset \vdash_{GL} \alpha.$$

## Modal system *K4LR*

- Axioms:
  - Tautologies;
  - Normality:  $\Box(\alpha \Rightarrow \beta) \Rightarrow (\Box\alpha \Rightarrow \Box\beta)$ ;
  - Transitivity axiom:  $\Box\alpha \Rightarrow \Box\Box\alpha$ ;
- Inference rules:
  - Modus Ponens: from  $\alpha$  and  $\alpha \Rightarrow \beta$  infer  $\beta$ ;
  - Necessitation: from  $\alpha$  infer  $\Box\alpha$ ;
  - Löb rule: from  $\Box\alpha \Rightarrow \alpha$  infer  $\alpha$ .

## Equivalence of *GL* and *K4LR*

$$\vdash_{GL} \alpha \text{ iff } \vdash_{K4LR} \alpha$$

The proof is straightforward in both directions.

## Kripke semantics

A model is a triple

$$M = \langle W, R, \Vdash \rangle$$

where:

- $W$  a non-empty set;
- $R$  is a transitive and conversely well-founded binary relation on  $W$ ;
- $\Vdash \subseteq W \times P$ .

Note: Since  $R$  is conversely well-founded, it is irreflexive.

## Local satisfaction

The local satisfaction relation at  $M$  is extended to  $L_{GL}$  as follows:

- $w \Vdash \neg\alpha$  if not  $w \Vdash \alpha$ ;
- $w \Vdash \alpha \Rightarrow \beta$  if not  $w \Vdash \alpha$  or  $w \Vdash \beta$ ;
- $w \Vdash \Box\alpha$  if  $w' \Vdash \alpha$  for every  $w'$  such that  $w R w'$ .

## Global satisfaction

$M \Vdash \alpha$  if  $w \Vdash \alpha$  for every  $w \in W$ .

## Kripke entailment

$\Gamma \vDash_{GL}^K \alpha$  if  $M \Vdash \alpha$  whenever  $M \Vdash \Gamma$ .

## Kripke validity

$\vDash_{GL}^K \alpha$  if  $\emptyset \vDash_{GL}^K \alpha$ .

## Kripke soundness

If  $\Gamma \vdash_{GL} \alpha$  then  $\Gamma \vDash_{GL}^K \alpha$ .

## Kripke weak completeness – Segerberg, 1971

If  $\Gamma \vDash_{GL}^K \alpha$  then  $\vdash_{GL} \alpha$ .

## Kripke non-compactness

Consider  $\Gamma = \{\diamond p_0\} \cup \{\Box(p_n \Rightarrow \diamond p_{n+1}) : n \in \mathbb{N}\}$ .

Then,  $\Gamma \vDash_{GL}^K \perp$  but no finite subset of  $\Gamma$  entails  $\perp$ .

## Kripke finite model property – Segerberg, 1971

If  $\alpha$  has a Kripke model then it has a finite Kripke model.

## Decidability

Theoremhood is decidable.

## Kripke weak completeness revised

*GL* is complete over the class of finite transitive irreflexive trees.

The latter result is useful in the proof of arithmetical completeness (Solovay, 1976).

Let  $P(\alpha)$  denote the set of propositional symbols occurring in  $\alpha$ .

A propositional symbol  $p$  is said to be *guarded* in  $\alpha$  if each of its occurrences in  $\alpha$  appears within the scope of a modality.

Let  $G(\alpha)$  denote the set of guarded propositional symbols occurring in  $\alpha$ .

Let  $\alpha_\beta^p$  be the formula obtained from  $\alpha$  by replacing every occurrence of  $p$  by  $\beta$ .

## Fixed point theorem – de Jongh / Sambin, 1975

Let  $p \in G(\alpha)$ . There is  $\beta$  such that:

- $P(\beta) \subseteq P(\alpha) \setminus \{p\}$ ;
- $\vdash_{GL} \beta \Leftrightarrow \alpha^p_\beta$ .

In these conditions  $\beta$  is said to be the fixed point of  $\lambda p.\alpha$ .

## Bernardi, 1976

Fixed points are unique up to equivalence.

## Smoryński, 1985 + Boolos, 1993

Algorithm for computing the fixed point for a guarded propositional symbol of a given formula.

### Example

The fixed point of  $\lambda p. \neg \Box p$  turns out to be  $\neg \Box \perp$ . Indeed:

$$\vdash_{GL} \neg \Box \perp \Leftrightarrow \neg \Box (\neg \Box \perp).$$

Arithmetically, the left to right implication states **Gödel's second incompleteness theorem**.

In the next section we shall make precise the relationship between *GL* and *PA*.

But, before, let us see if we could do away with modal logic...

## Idea

Encode modal logic in FOL as follows:

- take  $D$  to be  $W$ ;
- binary predicate  $r$  to be interpreted by  $R$ ;
- formulae translated as follows:
  - $\bar{p} = p(x)$  for each  $p \in P$ ;
  - $\overline{\neg\alpha} = \neg\bar{\alpha}$ ;
  - $\overline{\alpha \Rightarrow \beta} = \bar{\alpha} \Rightarrow \bar{\beta}$ ;
  - $\overline{\Box\alpha} = \forall y(r(x, y) \Rightarrow \bar{\alpha}_y^x)$  where  $y$  is chosen to be fresh in  $\bar{\alpha}$ .

## However

It does not work!

Indeed, it is not possible to axiomatize in FOL the envisaged theory. Namely, the converse well foundedness of  $R$  is not expressible in FOL.

It is rather curious that the provability of the FOL theory of arithmetic can be modally characterized (as we shall see next) but not in FOL.

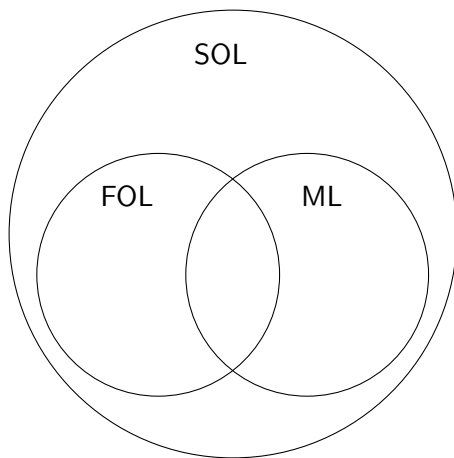
## Thomason, 1975

Modal logic corresponds to a fragment of SOL.

On the other hand, there are FOL properties of  $R$  that cannot be modally axiomatized. For instance, irreflexivity:  $\forall x \neg r(x, x)$ .

## Goldblatt-Thomason, 1974

Necessary and sufficient condition for a FOL property of  $R$  to be modally expressible.



# GL versus PA

## Arithmetical semantics

A realization is a map

$$f : L_{GL} \rightarrow L_{PA}$$

such that:

- $f(\perp) = \perp$ ;
- $f(\neg \alpha) = \neg f(\alpha)$ ;
- $f(\alpha \Rightarrow \beta) = f(\alpha) \Rightarrow f(\beta)$ ;
- $f(\Box \alpha) = \text{Prov}(\ulcorner f(\alpha) \urcorner)$ .

## Satisfaction

$f \Vdash \alpha$  if  $\vdash_{PA} f(\alpha)$ .

## Arithmetical validity

$\vDash_{GL}^A \alpha$  if  $f \Vdash \alpha$  for every realization  $f$ .

## Arithmetical weak soundness

If  $\vdash_{GL} \alpha$  then  $\models_{GL}^A \alpha$ .

Corollary of Löb's theorem and conditions (via *K4LR*).

## Arithmetical weak completeness – Solovay, 1976

If  $\models_{GL}^A \alpha$  then  $\vdash_{GL} \alpha$ .

Landmark result that capitalizes on the finite model property established by Segerberg for Kripke semantics of *GL*. Solovay devised a way of representing finite models in the language of *PA*.

## Formalized version of Löb's theorem

Theorem:  $\vdash_{PA} \text{Prov}(\ulcorner \text{Prov}(\ulcorner \alpha \urcorner) \Rightarrow \alpha \urcorner) \Rightarrow \text{Prov}(\ulcorner \alpha \urcorner)$

Immediate consequence of the arithmetical soundness of *GL*.

## What is the significance of $GL$ ?

- $GL$  captures everything that  $PA$  can say in modal terms about its own provability predicate:
- Solovay's proof of arithmetical weak completeness of  $GL$  provides an alternative way of finding valid arithmetical sentences that are not provable in  $PA$ .
- Furthermore, Solovay's theorem shows that what the undecidable theory  $PA$  can modally say about its own provability is captured by a decidable modal logic with a very simple possible world semantics.

## Provability logic today

How robust is Solovay's result?

That is, for which theories of arithmetic does it hold?

de Jongh, Jumelet and Montagna, 1991

Solovay's theorem holds for any sound theory  $T$  extending  $I\Delta_0 + EXP$ .

Can this be improved?

Answer seems to depend on difficult open problems in complexity theory.

What about *HA* (Heyting arithmetic)?

Visser, van Benthem, de Jongh and Renardel de Lavalette, 1995

Not known what should be the envisaged provability logic, but it should at least add the following principles to intuitionistic logic:

- $\Box(\Box\alpha \Rightarrow \alpha) \Rightarrow \Box\alpha$ ;
- $\Box\neg\neg\Box\alpha \Rightarrow \Box\Box\alpha$ ;
- $\Box(\neg\neg\Box\alpha \Rightarrow \Box\alpha) \Rightarrow \Box\Box\alpha$ ;
- $\Box(\alpha \vee \beta) \Rightarrow \Box(\alpha \vee \Box\beta)$ .

For more recent developments see Artemov and Iemhoff, 2006 (labelled calculus).

## Other developments

- Provability logic with propositional quantifiers (Shavrukov, 1997);
- Predicate provability logic – not axiomatizable (Vardanyan, 1986);
- Polymodal provability logic (Beklemishev, 1996);
- Interpretability logic (Berarducci, 1990);
- Applications of provability logic in proof theory (Beklemishev, 1999);
- Magari algebras (Shavrukov, 1993, 1997).

## Where to learn more...

Verbrugge, 2003

Provability logic, Stanford Encyclopedia of Philosophy, WWW.

Goldblatt, 2005

Mathematical modal logic: A view of its evolution, Handbook of the History of Logic, volume 6, Elsevier.

Japaridze and de Jongh, 1998

The logic of provability, Handbook of Proof Theory, Elsevier.

Boolos, 1993

The Logic of Provability, Cambridge UP.