Gödel – Löb Provability Logic

Logic II, University of Pennsylvania, Spring 2017

GL in a Hilbert-style Form

Language: $p, q, r, \ldots; \wedge, \vee, \rightarrow, \bot, \Box$.

Axioms:

- Classical propositional calculus.
- $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$ (normality).
- $\Box(\Box A \rightarrow A) \rightarrow \Box A$ (Löb).

Inference rules:

$$\frac{A \quad A \rightarrow B}{B}\ (\mathrm{MP}) \qquad \frac{A}{\Box A}\ (\mathrm{Nec})$$

GL in a Hilbert-style Form

Language: $p, q, r, \ldots;\ \wedge,\ \vee,\ \to,\ \bot,\ \Box$.

Axioms:

- Classical propositional calculus.
- $\Box(A \to B) \to (\Box A \to \Box B)$ (normality).
- $\Box(\Box A \to A) \to \Box A$ (Löb).

Inference rules:

$$\frac{A \quad A \to B}{B}\ (\mathrm{MP}) \qquad \frac{A}{\Box A}\ (\mathrm{Nec})$$

Lemma

GL $\vdash \Box A \to \Box\Box A$ *(transitivity).*

Kripke Semantics for **GL**

GL-frame: $\mathcal{F} = \langle W, R \rangle$, where

1. R is transitive;
2. there are no infinite chains $x_0 \, R \, x_1 \, R \, x_2 \, R \ldots$

Kripke Semantics for **GL**

GL-frame: $\mathcal{F} = \langle W, R \rangle$, where

1. R is transitive;
2. there are no infinite chains $x_0 \, R \, x_1 \, R \, x_2 \, R \ldots$

Finite **GL**-frames = transitive irreflexive frames.

Kripke Semantics for **GL**

GL-frame: $\mathcal{F} = \langle W, R \rangle$, where

1. R is transitive;
2. there are no infinite chains $x_0 \, R \, x_1 \, R \, x_2 \, R \ldots$

Finite **GL**-frames = transitive irreflexive frames.

GL-model: $\mathcal{M} = \langle \mathcal{F}, v \rangle$.

Kripke Semantics for **GL**

GL-frame: $\mathcal{F} = \langle W, R \rangle$, where

1. R is transitive;
2. there are no infinite chains $x_0 \, R \, x_1 \, R \, x_2 \, R \ldots$

Finite **GL**-frames = transitive irreflexive frames.

GL-model: $\mathcal{M} = \langle \mathcal{F}, v \rangle$.

Theorem

GL *is sound and complete:*

GL $= \mathrm{Log}($*all* **GL**-*frames*$) = \mathrm{Log}($*all finite* **GL**-*frames*$)$.

Sequent Calculus $\mathbf{GL}^G$

Tight negation: formulae are built from $\bot$, $\top$, p_1, $\bar{p}_1$, p_2, $\bar{p}_2$, … using $\vee$, $\wedge$, $\Box$, $\Diamond$ $(= \neg\Box\neg)$.
Negation $(\neg)$ and implication $(\rightarrow)$ are meta-operations.

One-side sequents: a sequent Γ is a finite set of formulae.
(Informal meaning: Γ means $\bigvee_{A \in \Gamma} A$.)

Axioms: $\top$ $\quad p, \bar{p}$

Inference rules:

$$\frac{\Gamma, A_1 \quad \Gamma, A_2}{\Gamma, A_1 \wedge A_2}\ (\wedge) \qquad \frac{\Gamma, A_i}{\Gamma, A_1 \vee A_2}\ (\vee) \qquad \frac{\Diamond\Gamma, \Gamma, \Diamond\neg A, A}{\Diamond\Gamma, \Box A}\ (\text{Löb})$$

$$\frac{\Gamma}{\Gamma, \Delta}\ (\mathrm{weak}) \qquad \frac{\Gamma, A \quad \Delta, \neg A}{\Gamma, \Delta}\ (\mathrm{cut})$$

Cut Elimination and Kripke Completeness

Theorem

T.F.A.E.:

1. $\mathbf{GL} \vdash A$;
2. $A \in \mathrm{Log}(\textit{all } \mathbf{GL}\textit{-frames})$;
3. $A \in \mathrm{Log}(\textit{all finite } \mathbf{GL}\textit{-frames})$;
4. $\mathbf{GL}^G \vdash \{A\}$ *without* (cut);
5. $\mathbf{GL}^G \vdash \{A\}$.

Cut Elimination and Kripke Completeness

Theorem

T.F.A.E.:

1. $\mathbf{GL} \vdash A$;
2. $A \in \mathrm{Log}(\textit{all } \mathbf{GL}\textit{-frames})$;
3. $A \in \mathrm{Log}(\textit{all finite } \mathbf{GL}\textit{-frames})$;
4. $\mathbf{GL}^G \vdash \{A\}$ *without* (cut);
5. $\mathbf{GL}^G \vdash \{A\}$.

The only non-trivial step is (3) $\Rightarrow$ (4).

Models from Sequents

Saturated sequent Γ:

1. if $A \vee B \in \Gamma$, then $A \in \Gamma$ and $B \in \Gamma$;
2. if $A \wedge B \in \Gamma$, then $A \in \Gamma$ or $B \in \Gamma$.

Models from Sequents

Saturated sequent Γ:

1. if $A \vee B \in \Gamma$, then $A \in \Gamma$ and $B \in \Gamma$;
2. if $A \wedge B \in \Gamma$, then $A \in \Gamma$ or $B \in \Gamma$.

Saturation lemma: if $\mathbf{GL}^G \nvdash \Gamma$ without (cut), then there exists a saturated $\Gamma' \supseteq \Gamma$ such that $\mathbf{GL}^G \nvdash \Gamma'$ and $\Gamma' \subseteq \mathrm{SubFm}(\Gamma)$.

Models from Sequents

Saturated sequent Γ:

1. if $A \vee B \in \Gamma$, then $A \in \Gamma$ and $B \in \Gamma$;
2. if $A \wedge B \in \Gamma$, then $A \in \Gamma$ or $B \in \Gamma$.

Saturation lemma: if $\mathbf{GL}^G \nvdash \Gamma$ without (cut), then there exists a saturated $\Gamma' \supseteq \Gamma$ such that $\mathbf{GL}^G \nvdash \Gamma'$ and $\Gamma' \subseteq \mathrm{SubFm}(\Gamma)$.

Counter-model.

- Let $\mathbf{GL}^G \nvdash \Gamma_0$ and let Φ be $\mathrm{SubFm}(\Gamma)$ closed under $\neg$.
- Let W be the set of all non-derivable saturated sequents $\Gamma' \subseteq \Phi$.
- $\Gamma \, R \, \Delta$, if
  1. if $\Diamond B \in \Gamma$, then $B, \Diamond B \in \Delta$;
  2. there exists $\Diamond B \in \Delta$ such that $\Diamond B \notin \Gamma$.

Models from Sequents

Saturated sequent Γ:

1. if $A \vee B \in \Gamma$, then $A \in \Gamma$ and $B \in \Gamma$;
2. if $A \wedge B \in \Gamma$, then $A \in \Gamma$ or $B \in \Gamma$.

Saturation lemma: if $\mathbf{GL}^G \nvdash \Gamma$ without (cut), then there exists a saturated $\Gamma' \supseteq \Gamma$ such that $\mathbf{GL}^G \nvdash \Gamma'$ and $\Gamma' \subseteq \mathrm{SubFm}(\Gamma)$.

Counter-model.

- Let $\mathbf{GL}^G \nvdash \Gamma_0$ and let Φ be $\mathrm{SubFm}(\Gamma)$ closed under $\neg$.
- Let W be the set of all non-derivable saturated sequents $\Gamma' \subseteq \Phi$.
- $\Gamma \, R \, \Delta$, if
  1. if $\Diamond B \in \Gamma$, then $B, \Diamond B \in \Delta$;
  2. there exists $\Diamond B \in \Delta$ such that $\Diamond B \notin \Gamma$.

$\langle W, R \rangle$ is a **GL**-frame.

Models from Sequents

Valuation: $v(\Gamma, p) = 1$ iff $p \notin \Gamma$.

$\mathcal{M} = \langle W, R, v \rangle$

Main Semantic Lemma. If $A \in \Gamma$, then $\mathcal{M}, \Gamma \nVdash A$.

Corollary: completeness. Saturate $\Gamma_0 \supseteq \Gamma_0'$, then $\mathcal{M}, \Gamma_0' \nVdash \bigvee \Gamma_0$.

Proof: induction on A.

Models from Sequents

Interesting induction cases:

- $A = \Diamond B \in \Gamma$.
  Then for any $\Delta \in R(\Gamma)$ we have $B \in \Delta$, therefore (IH) $\Delta \not\Vdash B$, therefore $\Gamma \not\Vdash \Diamond B$.

Models from Sequents

Interesting induction cases:

- $A = \Diamond B \in \Gamma$.
  Then for any $\Delta \in R(\Gamma)$ we have $B \in \Delta$, therefore (IH) $\Delta \not\Vdash B$, therefore $\Gamma \not\Vdash \Diamond B$.
- $A = \Box B \in \Gamma$.
  Let $\Delta = \{C, \Diamond C \mid \Diamond C \in \Gamma\} \cup \{\Diamond\neg B, B\}$.
  Claim: $\not\vdash \Delta$. Otherwise:

$$\dfrac{\dfrac{\Delta}{\{\Diamond C \mid C \in \Gamma\}, \Box B}\,(\text{Löb})}{\Gamma}\,(\text{weak})$$

Saturate $\Delta \subseteq \Delta'$. We have $\Gamma \, R \, \Delta'$ (irreflexivity: $\Diamond\neg B \in \Delta'$, $\notin \Gamma$).

Arithmetical Interpretation of **GL**

v : modal formulae $\to$ closed arithmetical formulae

- $v(p_i)$ arbitrary;
- $v(A \vee B) = v(A) \vee v(B)$, $v(A \wedge B) = v(A) \wedge v(B)$, $v(A \to B) = v(A) \to v(B)$, $v(\bot) = \bot$;
- $v(\Box A) = \mathrm{Pr}_{\mathbf{PA}}(\ulcorner A \urcorner) = \exists y\, \mathrm{Prf}_{\mathbf{PA}}(y, \ulcorner A \urcorner)$.

Arithmetical Interpretation of **GL**

v : modal formulae $\to$ closed arithmetical formulae

- $v(p_i)$ arbitrary;
- $v(A \vee B) = v(A) \vee v(B)$, $v(A \wedge B) = v(A) \wedge v(B)$, $v(A \to B) = v(A) \to v(B)$, $v(\bot) = \bot$;
- $v(\Box A) = \mathrm{Pr}_{\mathbf{PA}}(\ulcorner A \urcorner) = \exists y\, \mathrm{Prf}_{\mathbf{PA}}(y, \ulcorner A \urcorner)$.

Autoprovability logic: $\mathbb{PL}_{\mathbf{PA}}(\mathbf{PA}) = \{A \mid \mathbf{PA} \vdash v(A) \text{ for any } v\}$.

Arithmetical Interpretation of **GL**

v: modal formulae $\to$ closed arithmetical formulae

- $v(p_i)$ arbitrary;
- $v(A \vee B) = v(A) \vee v(B)$, $v(A \wedge B) = v(A) \wedge v(B)$,
  $v(A \to B) = v(A) \to v(B)$, $v(\bot) = \bot$;
- $v(\Box A) = \mathrm{Pr}_{\mathbf{PA}}(\ulcorner A \urcorner) = \exists y\, \mathrm{Prf}_{\mathbf{PA}}(y, \ulcorner A \urcorner)$.

Autoprovability logic: $\mathbb{PL}_{\mathbf{PA}}(\mathbf{PA}) = \{A \mid \mathbf{PA} \vdash v(A) \text{ for any } v\}$.

Theorem (Solovay)

$\mathbb{PL}_{\mathbf{PA}}(\mathbf{PA}) = \mathbf{GL}$.

Arithmetical Interpretation of **GL**

v : modal formulae $\to$ closed arithmetical formulae

- $v(p_i)$ arbitrary;
- $v(A \vee B) = v(A) \vee v(B)$, $v(A \wedge B) = v(A) \wedge v(B)$, $v(A \to B) = v(A) \to v(B)$, $v(\bot) = \bot$;
- $v(\Box A) = \mathrm{Pr}_{\mathbf{PA}}(\ulcorner A \urcorner) = \exists y\, \mathrm{Prf}_{\mathbf{PA}}(y, \ulcorner A \urcorner)$.

Autoprovability logic: $\mathbb{PL}_{\mathbf{PA}}(\mathbf{PA}) = \{A \mid \mathbf{PA} \vdash v(A) \text{ for any } v\}$.

Theorem (Solovay)

$\mathbb{PL}_{\mathbf{PA}}(\mathbf{PA}) = \mathbf{GL}$.

Correctness ($\supseteq$) is due to Gödel-II (formalised Löb's theorem and Hilbert – Bernays conditions).

The Refugee Function

$\mathbf{GL} \nvdash A$
$\mathcal{M}$ is the Kripke countermodel with $W = \{1, \ldots, n\}$
(1 is the root, $\mathcal{M}, 1 \nVdash A$).

The Refugee Function

GL $\nvdash A$
$\mathcal{M}$ is the Kripke countermodel with $W = \{1, \ldots, n\}$
(1 is the root, $\mathcal{M}, 1 \nVdash A$).
Add a new root 0, with all variables false.

The Refugee Function

$\mathbf{GL} \nvdash A$
$\mathcal{M}$ is the Kripke countermodel with $W = \{1, \ldots, n\}$
(1 is the root, $\mathcal{M}, 1 \nVdash A$).
Add a new root 0, with all variables false.

Refugee function:

$$h(0) = 0$$

$$h(t+1) = \begin{cases} z, & \text{if } h(t)\, R\, z \text{ and } \mathrm{Prf}_{\mathbf{PA}}(t+1, \ulcorner \neg S_z \urcorner) \\ h(t) & \text{otherwise} \end{cases}$$

$$S_x = \text{``}\lim_{t\to\infty} h(t) = x\text{''} = \exists t_0\, \forall t > t_0\; h(t) = x$$

(Defined using the fixpoint construction.)

The Refugee Function: Properties

- **PA** $\vdash t_1 < t_2 \rightarrow (h(t_1) \, R \, h(t_2) \lor h(t_1) = h(t_2))$

The Refugee Function: Properties

- **PA** $\vdash t_1 < t_2 \to (h(t_1)\, R\, h(t_2) \vee h(t_1) = h(t_2))$
- **PA** $\vdash \bigvee_{x \in W} S_x$

The Refugee Function: Properties

- **PA** $\vdash t_1 < t_2 \to (h(t_1)\, R\, h(t_2) \vee h(t_1) = h(t_2))$
- **PA** $\vdash \bigvee_{x \in W} S_x$
- **PA** $\vdash \neg(S_x \wedge S_y)$ for $x \neq y$

The Refugee Function: Properties

- **PA** $\vdash t_1 < t_2 \to (h(t_1)\, R\, h(t_2) \lor h(t_1) = h(t_2))$
- **PA** $\vdash \bigvee_{x \in W} S_x$
- **PA** $\vdash \neg(S_x \land S_y)$ for $x \neq y$
- **PA** $\vdash S_x \to \neg \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \neg S_y \urcorner)$, if $x\, R\, y$
  (otherwise there exists an arbitrarily long proof, serving as a visa to migrate to y)

The Refugee Function: Properties

- **PA** $\vdash t_1 < t_2 \to (h(t_1) \, R \, h(t_2) \lor h(t_1) = h(t_2))$
- **PA** $\vdash \bigvee_{x \in W} S_x$
- **PA** $\vdash \neg(S_x \land S_y)$ for $x \neq y$
- **PA** $\vdash S_x \to \neg \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \neg S_y \urcorner)$, if $x \, R \, y$
  (otherwise there exists an arbitrarily long proof, serving as a visa to migrate to y)
- If $x \neq 0$, then **PA** $\vdash S_x \to \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \bigvee_{y \in R(x)} S_y \urcorner)$.

The Refugee Function: Properties

- $\mathbf{PA} \vdash t_1 < t_2 \to (h(t_1)\, R\, h(t_2) \vee h(t_1) = h(t_2))$
- $\mathbf{PA} \vdash \bigvee_{x \in W} S_x$
- $\mathbf{PA} \vdash \neg(S_x \wedge S_y)$ for $x \neq y$
- $\mathbf{PA} \vdash S_x \to \neg \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \neg S_y \urcorner)$, if $x\, R\, y$
  (otherwise there exists an arbitrarily long proof, serving as a visa to migrate to y)
- If $x \neq 0$, then $\mathbf{PA} \vdash S_x \to \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \bigvee_{y \in R(x)} S_y \urcorner)$.

$$v(p) = \bigvee_{x \Vdash p} S_x.$$

The Refugee Function: Properties

- $\mathbf{PA} \vdash t_1 < t_2 \to (h(t_1)\, R\, h(t_2) \vee h(t_1) = h(t_2))$
- $\mathbf{PA} \vdash \bigvee_{x \in W} S_x$
- $\mathbf{PA} \vdash \neg(S_x \wedge S_y)$ for $x \neq y$
- $\mathbf{PA} \vdash S_x \to \neg \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \neg S_y \urcorner)$, if $x\, R\, y$
(otherwise there exists an arbitrarily long proof, serving as a visa to migrate to y)
- If $x \neq 0$, then $\mathbf{PA} \vdash S_x \to \mathrm{Pr}_{\mathbf{PA}}(\ulcorner \bigvee_{y \in R(x)} S_y \urcorner)$.

$$v(p) = \bigvee_{x \Vdash p} S_x.$$

Lemma

For every $x \neq 0$ and every modal formula A:

1. *if $x \Vdash A$, then $\mathbf{PA} \vdash S_x \to v(A)$;*
2. *if $x \nVdash A$, then $\mathbf{PA} \vdash S_x \to \neg v(A)$.*