On the Goldblatt-Thomason Theorem for Interpretability Logic

Fèlix Frigola González
Co-authored with Joost J. Joosten,
Vicent Navarro Arroyo and Cosimo Perini Brogi

Universitat de Barcelona

Logic Colloquium 2026

Provability vs. Interpretability

- Provability logic (**GL**)
  - $\Box A$ means that "A is provable".
  - Describes what a theory can prove about its own proofs.
  - **GL** works for sufficiently strong, sound theories.

- Interpretability logic (**IL**)
  - $A \rhd B$ means that "$T + A$ interprets $T + B$".
  - Describes relative strength between extensions of a theory.
  - Different theories (PA, GB, PRA, IΣ ...) give different logics **IL**(T).

The Logic IL

IL extends classical propositional logic by:

- L1 $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
- L2 $\Box\varphi \to \Box\Box\varphi$
- L3 $\Box(\Box\varphi \to \varphi) \to \Box\varphi$

- J1 $\Box(\alpha \to \beta) \to (\alpha \rhd \beta)$
- J2 $(\alpha \rhd \beta) \wedge (\beta \rhd \gamma) \to (\alpha \rhd \gamma)$
- J3 $(\alpha \rhd \gamma) \wedge (\beta \rhd \gamma) \to (\alpha \vee \beta) \rhd \gamma$
- J4 $(\alpha \rhd \beta) \to (\Diamond\alpha \to \Diamond\beta)$
- J5 $\Diamond\alpha \rhd \alpha$

together with Modus Ponens and Necessitation Rule:

$$\text{If } \vdash_{\mathsf{IL}} \varphi, \text{ then } \vdash_{\mathsf{IL}} \Box\varphi. \qquad \text{(Nec.)}$$

Veltman Frames

A **Veltman frame** is a structure

$$\mathcal{F} = \langle W, R, \{S_x\}_{x \in W} \rangle$$

such that

- R is transitive;
- R is conversely well-founded;
- every S_x is reflexive;
- every S_x is transitive;
- $R \cap R[x] \subseteq S_x$.

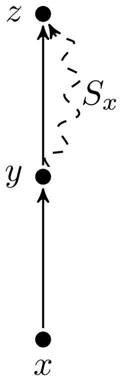


Figure: $R \cap R[x] \subseteq S_x$.

Interpretability requires a family of relations $\{S_w\}_{w \in W}$ rather than a single accessibility relation.

Veltman Frames

$$x \Vdash \varphi \rhd \psi :\iff \forall y\Big(xRy \wedge y \Vdash \varphi \Rightarrow \exists z\,(yS_x z \wedge z \Vdash \psi)\Big).$$

This additional structure makes correspondence theory considerably harder than in ordinary modal logic.

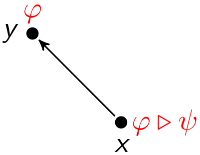
$\bullet\, \varphi \rhd \psi$

x

Veltman Frames

$$x \Vdash \varphi \rhd \psi :\iff \forall y\Big(xRy \wedge y \Vdash \varphi \Rightarrow \exists z\,(yS_xz \wedge z \Vdash \psi)\Big).$$

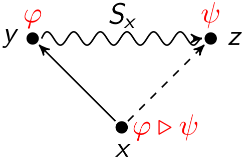
This additional structure makes correspondence theory considerably harder than in ordinary modal logic.



Veltman Frames

$$x \Vdash \varphi \rhd \psi : \iff$$
$$\forall y \Big(xRy \wedge y \Vdash \varphi \Rightarrow \exists z \, (yS_x z \wedge z \Vdash \psi) \Big).$$

This additional structure makes correspondence theory considerably harder than in ordinary modal logic.



ILM is the logic of theries with full induction.

ILM: (Montagna's principle)

$$\varphi \rhd \psi \to (\varphi \wedge \Box\phi) \rhd (\psi \wedge \Box\phi)$$

whose **frame condition** is:

$$xRyS_xzRu \to yRu.$$

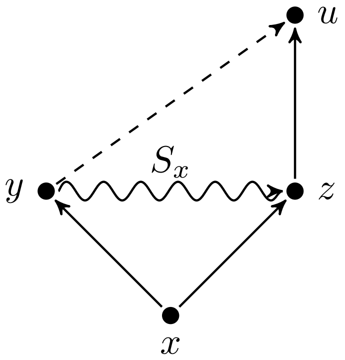


Figure: ILM frame condition.

Frame Definability

A class of frames K is **modally definable** if:

$$\exists \Gamma \quad \forall F \, (F \models \Gamma \iff F \in \mathsf{K})$$

If a class of Frames is modally definable we can characterize it by an axiom/aixom scheme.

Open Question

What is the interpretability logic of all reasonable arithmetical theories? (IL(All))

The Goldblatt-Thomason Theorem

Theorem (Goldblatt-Thomason)

Let K *be an elementary class of Kripke Frames.* K *is modally definable if and only if* K *is closed under Disjoint Unions, Generated subframes, Bounded Morphic Images and reflecting Ultrafilter Extensions*

Frames are elementary if they can be defined by a set of first-order formulae.
All Veltman Frames are Conversely Well-Founded, which can only be characterized at least by a second order formula. Therefore, infinite Veltman Frames are **not elementary**.

Goldblatt-Thomason: Finite Frames

For now we will restrict ourselves to finite Veltman Frames:

Theorem (FfG, JjJ, VnA, CpB)

Let K *be class of* **finite** *Veltman Frames.* K *is modally definable if and only if* K *is closed under Disjoint Unions, Generated subframes, Bounded Morphic Images.*

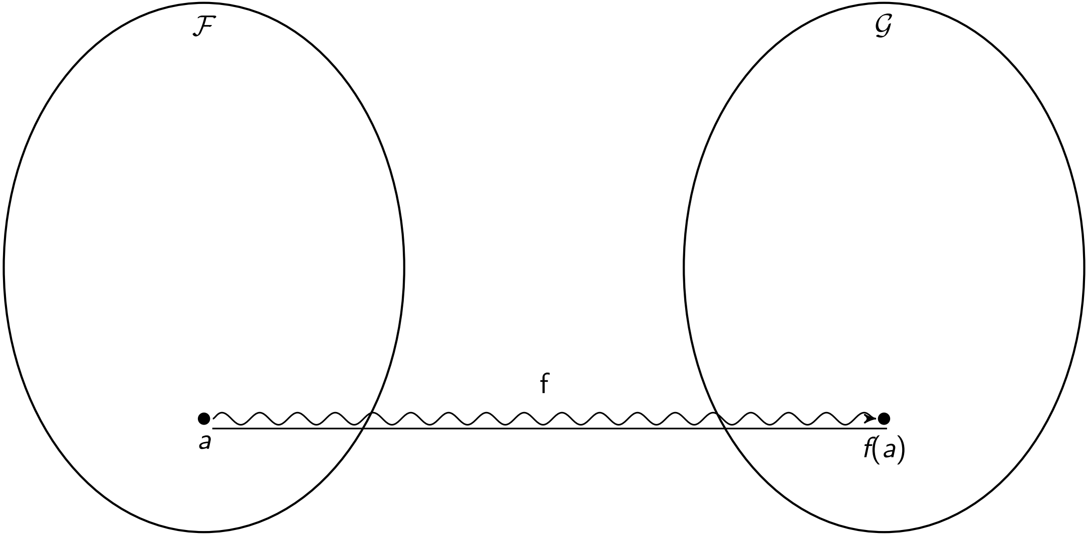
Definition of Bounded Morphism

Definition

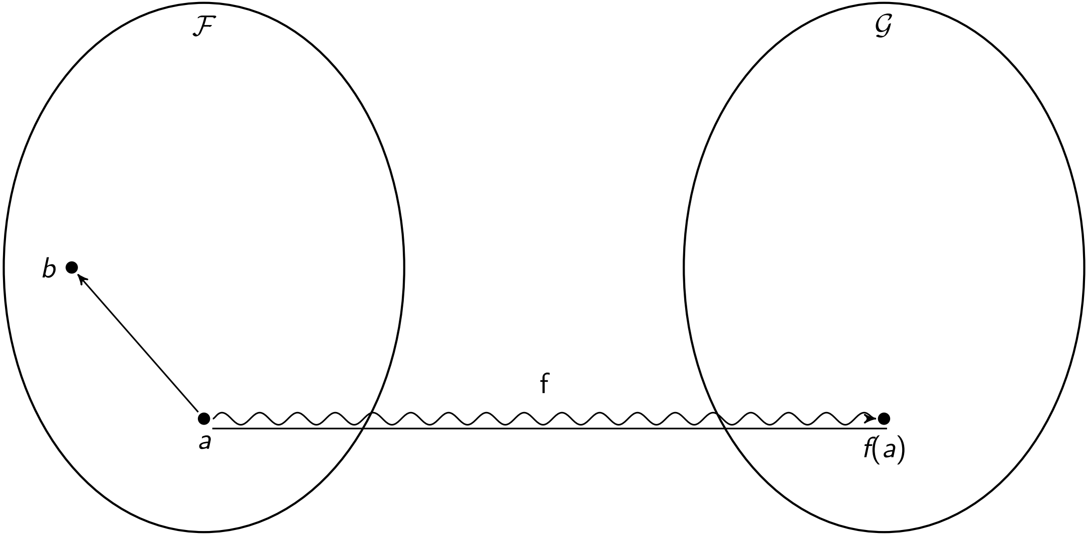
Given $F = \langle W, R, \{S_w\}_{w \in W} \rangle$ and $F' = \langle W', R', \{S'_u\}_{u \in W'} \rangle$ Veltman Frames we say that F' is a bounded morphic image of F via f whenever:

1. f is surjective.
2. wRv then $f(w)R'f(v)$ and for all u', if $f(v)S'_{f(w)}u'$ then there is some u such that vS_wu and $f(u) = u'$.
3. $f(w)R'v'$ then there is some v such that wRv and $f(v) = v'$ and for all u, if vS_wu then $f(v)S_{f(w)}f(u)$.

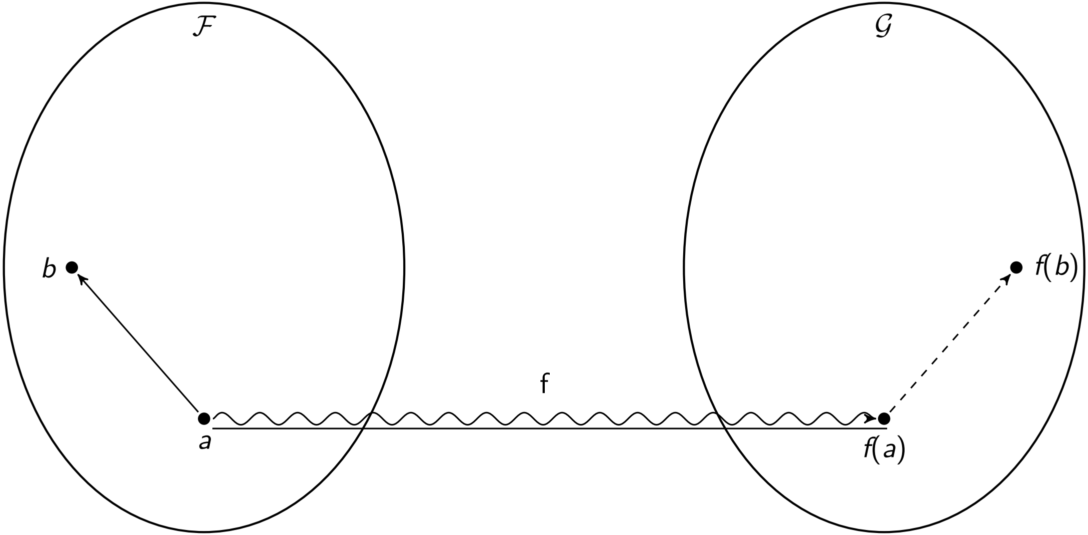
Bounded Morphisms



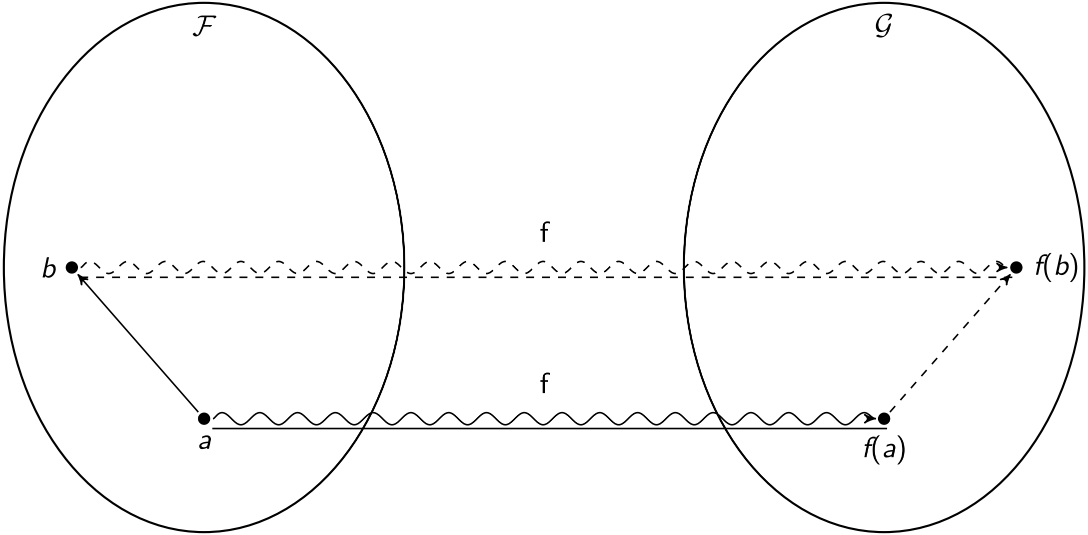
Bounded Morphisms



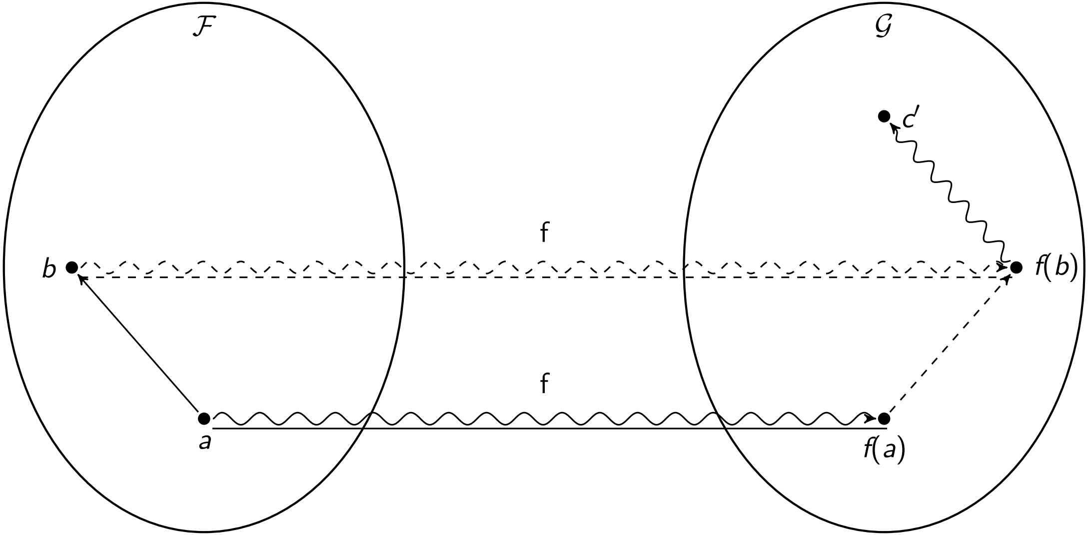
Bounded Morphisms



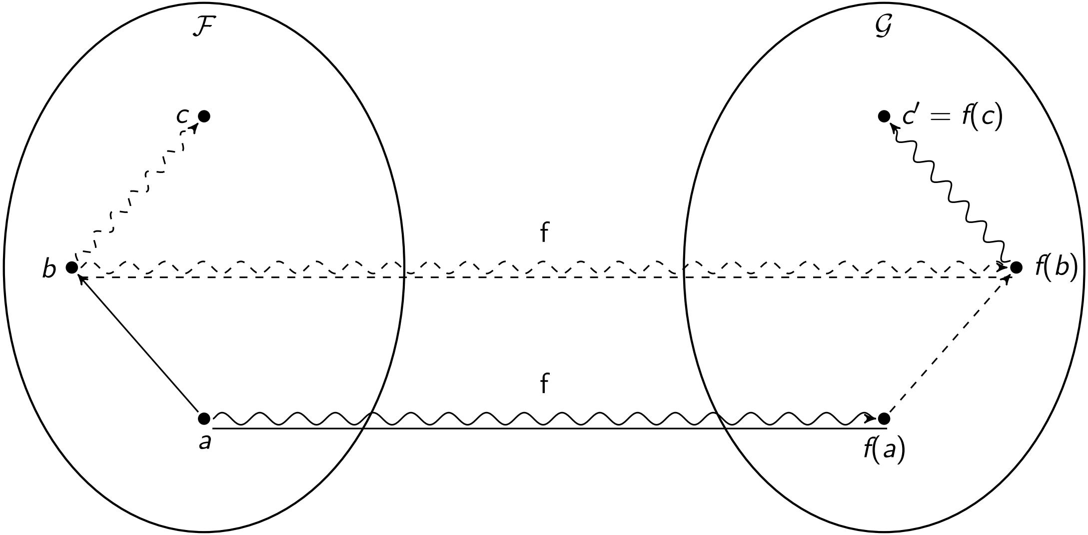
Bounded Morphisms



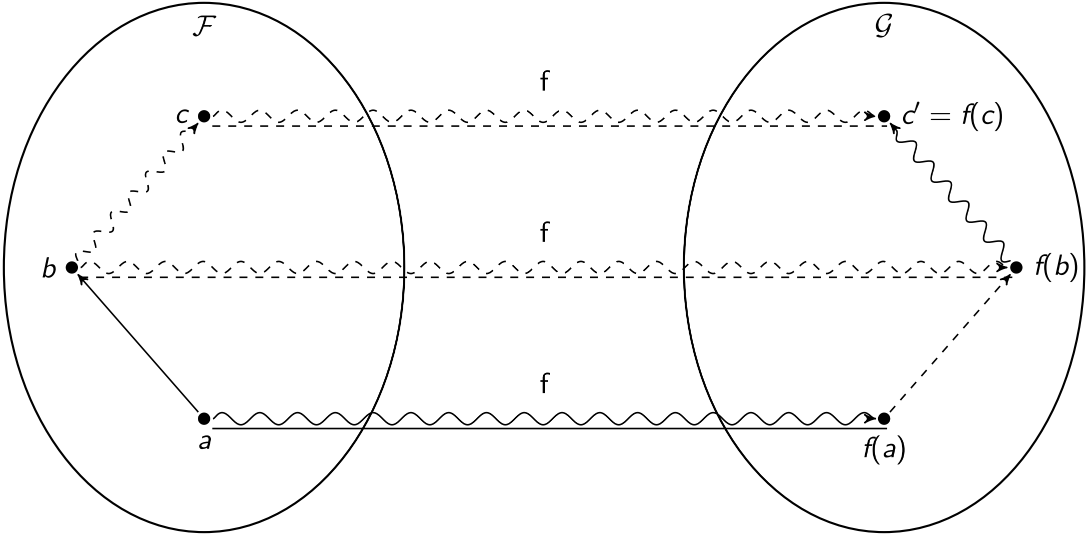
Bounded Morphisms



Bounded Morphisms



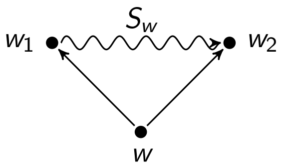
Bounded Morphisms



The Finite Goldblatt-Thomason Theorem

The Golblatt-Thomason Theorem for finite Veltman Frames is proven via formulas. Given a point-generated Veltman Frame F_w we define $\varphi_{F,w}$ the Jankov-Fine fro F_w as follows: Let W_i be an enumeration on W such that $w_0 = w$.
We use $\boxdot\varphi$ to denote $\varphi \wedge \Box\varphi$.

1. p_0
2. $\boxdot(p_i \to \neg p_j)$ if $i \neq j$
3. $\boxdot(p_i \to \Diamond p_j)$ if $w_i R w_j$
4. $\boxdot(p_i \to \neg\Diamond p_j)$ if $\neg w_i R w_j$
5. $\boxdot(p_i \to (p_j \rhd p_k))$ if $w_i R w_j$ and $w_j S_{w_i} w_k$
6. $\boxdot(p_i \to (p_j \rhd \neg p_k))$ if $w_i R w_j$ and $\neg w_j S_{w_i} w_k$



Jankov-Fine Formulas

Lemma

Let $F_w = \langle W, R, \{S_x\}_{x \in W} \rangle$ *be a point-generated finite Veltman Frame and let* w *be its root. Let* $\varphi_{F,w}$ *be the corresponding Jankov-Fine formula for* F *and* w*. Then, for all Veltman Frames* G

$$\exists V, \exists v (\langle G, V \rangle, v \Vdash \varphi_{F.w}) \iff \exists f \colon G_v \twoheadrightarrow F_w \,(f \textit{ is a bounded morphism})$$

Proof idea:

- Let K be a class of finite Veltman Frames closed under $\uplus$, F_w, BMI.
- Let $Th(\mathrm{K})$ be the Theory of $\mathrm{K}(\{\varphi : \forall F \in \mathrm{K}\,(F \models \varphi)\})$.
  We want to see that for all frames F, $F \models Th(\mathrm{K}) \iff F \in \mathrm{K}$.
- Let $F \models Th(\mathrm{K})$. Assume $F = F_w$, then $\varphi_{F,w}$ is satisfiable in F.
- $\neg\varphi_{F,w} \notin Th(\mathrm{K})$. Thus for some $G \in \mathrm{K}$ G satisfies $\varphi_{F,w}$.

$$\begin{tikzcd} G_v \arrow[r, dash] \arrow[d, dash] & F_w \arrow[d, dash] \\ G & F \end{tikzcd}$$

Conclusion

- The Goldblatt-Thomason Theorem for Finite Veltman Frames holds!
- Ultrafilter Extensions are defined and the main lemmas hold.
- If a class K of Veltman Frames is definable it is closed under disjoint unions, generated subframes, bounded morphic images and reflecting ultrafilter extensions.
- Only the issue of Elementarity remains open.

Thank You!