Results and ideas
on proof theory for interpretability logics

Cosimo Perini Brogi

University of Barcelona

The Proof Society Workshop
November 11-12, 2022

Interpretability logics

(Visser 1988,1990), (de Jongh and Veltman 1990), $\cdots$

$A \rhd B$	$\longmapsto$	$Int_{\mathrm{T}}(\ulcorner A^* \urcorner, \ulcorner B^* \urcorner)$	in *interpretability logics*
	$\approx$		
$\Box A$	$\longmapsto$	$Bew_T(\ulcorner A^* \urcorner)$	in *provability logic*

- Axiom schemas of $\mathbb{CPC}$;
- schema IL2 : $(A \rhd B) \to (B \rhd C) \to (A \rhd C)$;
- schema IL3 : $(A \rhd C) \to (B \rhd C) \to (A \vee B \rhd C)$;
- schema IL-Löb: $A \rhd (A \wedge (A \rhd \bot))$;
- MP Rule $\dfrac{A \to B \qquad A}{B}$;
- $\rhd$Rule $\dfrac{A \to B}{A \rhd B}$.

We define

$$\Box A := \neg A \rhd \bot, \text{ and } \Diamond A := \neg \Box \neg A.$$

Let us define as proper extensions of $\mathbb{IL}$

- $\mathbb{ILM} := \mathbb{IL} + \mathsf{M}$, where
$$\mathsf{M} := (A \rhd B) \to ((A \wedge \Box C) \rhd (B \wedge \Box C))$$
is called the Montagna schema;
- $\mathbb{ILP} := \mathbb{IL} + \mathsf{P}$, where
$$\mathsf{P} := (A \rhd B) \to \Box(A \rhd B)$$
is called the persistence schema;
- $\mathbb{ILW} := \mathbb{IL} + \mathsf{W}$, where
$$\mathsf{W} := (A \rhd B) \to (A \rhd (B \wedge \Box \neg A))$$
is called the de Jongh-Visser schema;
- $\mathbb{ILKM}1 := \mathbb{IL} + \mathsf{KM1}$, where
$$\mathsf{KM1} := (A \rhd \Diamond \top) \to (\top \rhd \neg A);$$
- $\mathbb{ILM}_0 := \mathbb{IL} + \mathsf{M}_0$, where
$$\mathsf{M}_0 := (A \rhd B) \to ((\Diamond A \wedge \Box C) \rhd (B \wedge \Box C));$$

Each of these extensions can be characterised in terms of GVS semantics by imposing specific conditions to frames.

Interpretability logics

(Verbrugge 1992) semantics

A generalised Veltman frame $\mathcal{F}$ consists of

- a finite set $W \neq \varnothing$;
- a binary relation $R \subseteq W \times W$ which is irreflexive and transitive;
- a W-indexed set of relations $S_x \subseteq R[x] \times (\wp(R[x]) \setminus \{\varnothing\})$
  – where $R[x]$ is the set of R-accessible worlds from x;

satisfying the following conditions:

- Quasi-reflexivity: if xRy then $yS_x\{y\}$;
- Definiteness: if $xRyRz$ then $yS_x\{z\}$;
- Monotonicity: if yS_xa and $a \subseteq b \subseteq R[x]$ then yS_xb;
- Quasi-transitivity: if yS_xa and vS_xb_v for all $v \in a$, then $yS_x(\bigcup_{v\in a} b_v)$.

$x \Vdash A \rhd B$ iff for all y if xRy and $y \Vdash A$, then there exists an a such that yS_xa and $a \Vdash^{\forall} B$,

– where $a \Vdash^{\forall} B$ abbreviates the expression "for any $z \in a, z \Vdash B$".

Interpretability logics

(Verbrugge 1992) semantics

A generalised Veltman frame $\mathcal{F}$ consists of

- a finite set $W \neq \varnothing$;
- a binary relation $R \subseteq W \times W$ which is irreflexive and transitive;
- a W-indexed set of relations $S_x \subseteq R[x] \times (\wp(R[x]) \setminus \{\varnothing\})$
  – where $R[x]$ is the set of R-accessible worlds from x;

satisfying the following conditions:

- Quasi-reflexivity: if xRy then $yS_x\{y\}$;
- Definiteness: if $xRyRz$ then $yS_x\{z\}$;
- Monotonicity: if yS_xa and $a \subseteq b \subseteq R[x]$ then yS_xb;
- Quasi-transitivity: if yS_xa and vS_xb_v for all $v \in a$, then $yS_x(\bigcup_{v\in a} b_v)$.

$x \Vdash A \rhd B$ iff for all y if xRy and $y \Vdash A$, then there exists an a such that yS_xa and $a \Vdash^{\forall} B$,

– where $a \Vdash^{\forall} B$ abbreviates the expression "for any $z \in a, z \Vdash B$".

Verbrugge semantics is almost a geometric theory: each of its axioms has shape

$$\forall \vec{x}, \phi \to \psi$$

– where ϕ, ψ are FO formulas that do not contain $\forall$ or $\to$.

Quasi-transitivity and finiteness are an exception.
However,

- *finiteness is not a real issue* here; and
- there exist several variants of quasi-transitivity, including

$$\text{if } yS_x a \text{ and } z \in a \text{ and } zS_x b\text{, then } yS_x b,$$

which is geometric.

Therefore, it should be possible to formalise Verbrugge semantics into a sequent system.

Verbrugge semantics is almost a geometric theory: each of its axioms has shape

$$\forall \vec{x}, \phi \to \psi$$

– where ϕ, ψ are FO formulas that do not contain $\forall$ or $\to$.

Quasi-transitivity and finiteness are an exception.

However,

- *finiteness is not a real issue* here; and
- there exist several variants of quasi-transitivity, including

$$\text{if } yS_x a \text{ and } z \in a \text{ and } zS_x b\text{, then } yS_x b,$$

which is geometric.

Therefore, it should be possible to formalise Verbrugge semantics into a sequent system.

Design of good calculi

Formalising Verbrugge semantics

Verbrugge semantics is almost a geometric theory: each of its axioms has shape

$$\forall \vec{x}, \phi \rightarrow \psi$$

– where ϕ, ψ are FO formulas that do not contain $\forall$ or $\rightarrow$.

Quasi-transitivity and finiteness are an exception.
However,

- *finiteness is not a real issue* here; and
- there exist several variants of quasi-transitivity, including

  if $yS_x a$ and $z \in a$ and $zS_x b$, then $yS_x b$,

  which is geometric.

Therefore,[1] it should be possible to formalise Verbrugge semantics into a sequent system.

[1] After (Negri and von Plato 2001).

Design of good calculi

Formalised semantic reasoning

(Hakoniemi and Joosten 2016) designed labelled tableaux – based on standard Veltman semantics – for the basic system and some extensions; (Sasaki 2001) provided a cut free standard sequent calculus for IL.

Here I propose a modular family of sequent calculi for IL and its extensions.

The general idea is to *explicitly* internalise GVS in the G3-paradigm, following the well-established of *labelling*.

Design of good calculi

(Hakoniemi and Joosten 2016) designed labelled tableaux – based on standard Veltman semantics – for the basic system and some extensions; (Sasaki 2001) provided a cut free standard sequent calculus for IL.

Here I propose a modular family of sequent calculi for IL and its extensions.

The general idea is to *explicitly* internalise GVS in the G3-paradigm, following the well-established of *labelling*.

Labelled sequent calculi for interpretability

Starting point

(♯) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$,
then there exists an a such that $yS_x a$ and $a \Vdash^{\forall} B$,

Labelled sequent calculi for interpretability

Starting point

$(\sharp)$ $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$,
then there exists an a such that $yS_x a$ and $a \Vdash^{\forall} B$,

$(\sharp\flat)$ $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$,
then $y \Vdash \langle]_x B$.

Therefore

$$x \Vdash A \rhd B \quad \text{iff} \quad x \Vdash \Box(A \to \langle]_x B).$$

Moreover, in any irreflexive transitive finite frame

$$x \Vdash \Box A \quad \text{iff} \quad \text{for any } y,\ \text{if } xRy \text{ and } y \Vdash \Box A,\ \text{then } y \Vdash A.$$

Henceforth

$(\natural)$ $(x,x) \Vdash A \rhd B$ iff for all y, if xRy and $(y,x) \Vdash A \rhd B$,
then, if $y \Vdash A$, $y \Vdash \langle]_x B$.

Labelled sequent calculi for interpretability

(♯) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$, then there exists an a such that $yS_x a$ and $a \Vdash^{\forall} B$,

(♯♭) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$, then $y \Vdash \langle]_x B$.

Therefore

$$x \Vdash A \rhd B \quad \text{iff} \quad x \Vdash \Box(A \to \langle]_x B).$$

Moreover, in any irreflexive transitive finite frame

$$x \Vdash \Box A \quad \text{iff} \quad \text{for any } y, \text{ if } xRy \text{ and } y \Vdash \Box A, \text{ then } y \Vdash A.$$

Henceforth

(♮) $(x, x) \Vdash A \rhd B$ iff for all y, if xRy and $(y, x) \Vdash A \rhd B$, then, if $y \Vdash A$, $y \Vdash \langle]_x B$.

Labelled sequent calculi for interpretability

Starting point

(♯) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$, then there exists an a such that $yS_x a$ and $a \Vdash^{\forall} B$,

(♯♭) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$, then $y \Vdash \langle]_x B$.

Therefore

$$x \Vdash A \rhd B \quad \text{iff} \quad x \Vdash \Box(A \to \langle]_x B).$$

Moreover, in any irreflexive transitive finite frame

$$x \Vdash \Box A \quad \text{iff} \quad \text{for any } y, \text{ if } xRy \text{ and } y \Vdash \Box A, \text{ then } y \Vdash A.$$

Henceforth

(♮) $(x, x) \Vdash A \rhd B$ iff for all y, if xRy and $(y, x) \Vdash A \rhd B$, then, if $y \Vdash A$, $y \Vdash \langle]_x B$.

Labelled sequent calculi for interpretability

($\sharp$) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$,
then there exists an a such that $yS_x a$ and $a \Vdash^{\forall} B$,

($\sharp\flat$) $x \Vdash A \rhd B$ iff for all y, if xRy and $y \Vdash A$,
then $y \Vdash \langle]_x B$.

Therefore

$$x \Vdash A \rhd B \quad \text{iff} \quad x \Vdash \Box(A \to \langle]_x B).$$

Moreover, in any irreflexive transitive finite frame

$$x \Vdash \Box A \quad \text{iff} \quad \text{for any } y, \text{ if } xRy \text{ and } y \Vdash \Box A, \text{ then } y \Vdash A.$$

Henceforth

($\natural$) $(x, x) \Vdash A \rhd B$ iff for all y, if xRy and $(y, x) \Vdash A \rhd B$,
then, if $y \Vdash A$, $y \Vdash \langle]_x B$.

Initial sequents

$x : p, \Gamma \Rightarrow \Delta, x : p$

$(x, w) : A \rhd B, \Gamma \Rightarrow \Delta, (x, w) : A \rhd B$

Classical propositional rules: the usual ones

Local forcing rules

$$\frac{x : A, x \in A, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}{x \in A, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}\ \mathcal{L}\Vdash^{\forall} \qquad \frac{x \in a, \Gamma \Rightarrow \Delta, x : A}{\Gamma \Rightarrow \Delta, a \Vdash^{\forall} A}\ \mathcal{R}\Vdash^{\forall}_{(x!)}$$

Intermediate modality rules

$$\frac{yS_x a, a \Vdash^{\forall} A, \Gamma \Rightarrow \Delta}{y : \langle]_x A, \Gamma \Rightarrow \Delta}\ \mathcal{L}\langle]_{(a!)} \qquad \frac{yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A, a \Vdash^{\forall} A}{yS_x a, \Gamma \Rightarrow \Delta, y : \langle]_x A}\ \mathcal{R}\langle]$$

G3IL*

Core system

Interpretability modality rules

$$\dfrac{y \in R[x], (x,w) : A \rhd B, \Gamma \Rightarrow \Delta, y : A \qquad y : \langle]_w B, y \in R[x], (x,w) : A \rhd B, \Gamma \Rightarrow \Delta \qquad y \in R[x], (x,w) : A \rhd B, \Gamma \Rightarrow \Delta, (y,w) : A \rhd B}{y \in R[x], (x,w) : A \rhd B, \Gamma \Rightarrow \Delta} \; \mathcal{L}\rhd$$

$$\dfrac{y \in R[x], y : A, \Gamma, (y,w) : A \rhd B \Rightarrow \Delta, y : \langle]_w B}{\Gamma \Rightarrow \Delta, (x,w) : A \rhd B} \; \mathcal{R}\rhd_{(y!)}$$

$(x, w) \Vdash A \rhd B$ iff for all y, if xRy and $(y, w) \Vdash A \rhd B$, then, if $y \Vdash A$, $y \Vdash \langle]_w B$.

$G3IL^{\star}$
Core system

Rules for GVS

$$\frac{a \subseteq a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\; Refl\subseteq \qquad \frac{a \subseteq c, a \subseteq b, b \subseteq c, \Gamma \Rightarrow \Delta}{a \subseteq b, b \subseteq c, \Gamma \Rightarrow \Delta}\; Trans\subseteq$$

$$\frac{x \in b, x \in a, a \subseteq b, \Gamma \Rightarrow \Delta}{x \in a, a \subseteq b, \Gamma \Rightarrow \Delta}\; \mathcal{L}\subseteq$$

$$\frac{x \in \{x\}, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\; Sing$$

$$\frac{Atm(y), Atm(x), y \in \{x\}, \Gamma \Rightarrow \Delta}{Atm(x), y \in \{x\}, \Gamma \Rightarrow \Delta}\; Repl_1 \qquad \frac{Atm(x), Atm(y), y \in \{x\}, \Gamma \Rightarrow \Delta}{Atm(y), y \in \{x\}, \Gamma \Rightarrow \Delta}\; Repl_2$$

where $Atm(x)$ has one of the following forms: $x : p$, $x \in a$, $x \in \{z\}$, $x \in R[z]$, $z \in R[x]$, $xS_z a$, $zS_x a$.

$$\frac{}{x \in R[x], \Gamma \Rightarrow \Delta}\; Irrefl \qquad \frac{z \in R[x], y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}\; Trans$$

$$\frac{z \in a, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\; NE_{(z!)} \qquad \frac{y \in R[x], a \subseteq R[x], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\; DefS1$$

$$\frac{yS_x \{z\}, y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], \Gamma \Rightarrow \Delta}\; DefS2 \qquad \frac{yS_x b, yS_x a, a \subseteq b, b \subseteq R[x], \Gamma \Rightarrow \Delta}{yS_x a, a \subseteq b, b \subseteq R[x], \Gamma \Rightarrow \Delta}\; Mono$$

$$\frac{yS_x \{y\}, y \in R[x], \Gamma \Rightarrow \Delta}{y \in R[x], \Gamma \Rightarrow \Delta}\; Qrefl \qquad \frac{yS_x b, yS_x a, z \in a, zS_x b, \Gamma \Rightarrow \Delta}{yS_x a, z \in a, zS_x b, \Gamma \Rightarrow \Delta}\; Qtrans6$$

G3IL*

Rules for extensions

Additional rules for GVS

$$\frac{x \in a, y \in R[x], y \in R[a], \Gamma \Rightarrow \Delta}{y \in R[a], \Gamma \Rightarrow \Delta}\ Rset1_{(x!)} \qquad \frac{y \in R[a], x \in a, y \in R[x], \Gamma \Rightarrow \Delta}{x \in a, y \in R[x], \Gamma \Rightarrow \Delta}\ Rset2$$

$$\frac{yS_x a, y \in S_x^{-1}a, \Gamma \Rightarrow \Delta}{y \in S_x^{-1}a, \Gamma \Rightarrow \Delta}\ Sset1 \qquad \frac{y \in S_x^{-1}a, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ Sset2$$

$$\frac{c \subseteq a, c \subseteq b, c \subseteq a \cap b, \Gamma \Rightarrow \Delta}{c \subseteq a \cap b, \Gamma \Rightarrow \Delta}\ \cap_1 \qquad \frac{c \subseteq a \cap b, c \subseteq a, c \subseteq b, \Gamma \Rightarrow \Delta}{c \subseteq a, c \subseteq b, \Gamma \Rightarrow \Delta}\ \cap_2$$

$$\frac{}{x \in \varnothing, \Gamma \Rightarrow \Delta}\ \mathcal{L}\varnothing$$

Rules for interpretability principles – via semantics characterisation by (Verbrugge 1992), (Vuković 1999)

$$\frac{b \subseteq a, yS_x b, R[b] \subseteq R[y], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ M_{(b!)} \qquad \frac{z \in a, R_z \subseteq R[y], yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ KM1_{(z!)}$$

$$\frac{b \subseteq a, zS_y b, y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}\ P_{(b!)} \qquad \frac{b \subseteq a, yS_x b, R[b] \cap S_x^{-1}a \subseteq \varnothing, yS_x a, \Gamma \Rightarrow \Delta}{yS_x a, \Gamma \Rightarrow \Delta}\ W_{(b!)}$$

$$\frac{b \subseteq a, yS_x b, R[b] \subseteq R[y], y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}{y \in R[x], z \in R[y], zS_x a, \Gamma \Rightarrow \Delta}\ M_{0(b!)}$$

Theorem (PB 2022)

Any calculus in the family G3IL* satisfies the following properties:

- *Generalised initial sequents are derivable;*
- *Substitution rules for worlds and neighbourhoods are height-preserving admissible;*
- *Weakening rules are height preserving admissible;*
- *All the rules are invertible;*
- *Contraction rules are admissible;*
- *Cut is admissible.*

Some care is needed for proving *cut elimination*:
We had to generalise the strategy by (Negri 2005), and proceed by *ternary* transfinite induction – main induction on the size of the cut formula, secondary induction on the range of the cut label and tertiary induction on the cut height.

Theorem (PB 2022)

Any calculus in the family G3IL* *satisfies the following properties:*

- *Generalised initial sequents are derivable;*
- *Substitution rules for worlds and neighbourhoods are height-preserving admissible;*
- *Weakening rules are height preserving admissible;*
- *All the rules are invertible;*
- *Contraction rules are admissible;*
- *Cut is admissible.*

Some care is needed for proving *cut elimination*:
We had to generalise the strategy by (Negri 2005), and proceed by *ternary* transfinite induction – main induction on the size of the cut formula, secondary induction on the range of the cut label and tertiary induction on the cut height.

G3IL*

Semantic completeness

Each calculus in the family of G3IL* is *sound and complete* w.r.t. the appropriate class of Verbrugge frames.
This is shown by interpreting derivations in frames – soundness – and, indirectly, by proving the interpretability principles of each axiomatic calculus – completeness.

Conjecture

There exists a strategy making proof search in $\mathsf{G3KIL}^\star$ for a sequent of the form $\Rightarrow x : A$ always terminate in a finite number of steps. Moreover, from a failed proof search, it is possible to extract a countermodel to A belonging to appropriate class of generalised Veltman frames.[a]

[a] Already proven for the flattened language.

- A direct proof of completeness, via Schütte-Takeuti-Tait extraction of a countermodel;
- A certified theorem prover for IL and its extensions;
- Considering further systems, e.g. ILP_0 (not hard), ILR (not easy), ILF (not known).

Future work

Conjecture

There exists a strategy making proof search in G3KIL* for a sequent of the form $\Rightarrow x : A$ always terminate in a finite number of steps. Moreover, from a failed proof search, it is possible to extract a countermodel to A belonging to appropriate class of generalised Veltman frames.[a]

[a] Already proven for the flattened language.

- A direct proof of completeness, via Schütte-Takeuti-Tait extraction of a countermodel;
- A certified theorem prover for IL and its extensions;
- Considering further systems, e.g. ILP_0 (not hard), ILR (not easy), ILF (not known).

Many thanks for your attention!

Picture: *Geometric and wavy lines* by Myriam Thyes, 2014,