Algebraic semantics for interpretability logics

Teo Šestak

University of Zagreb
teo.sestak@fsb.unizg.hr

Logic and Applications 2025

Table of Contents

Table of Contents

Introduction

Algebraic semantics (for normal modal logics):

- boolean algebras with additional operators,
- robustly complete.

Interpretability logics:

- extension of provability logic **GL**,
- interpreted on Kripke-like frames called Veltman frames.

Goals of this talk:

- define algebras for interpretability logics,
- check the similarity between the properties of algebras for modal and interpretability logics.

Boolean algebras with operators

- boolean algebra with operators (BAO): a boolean algebra together with an operator $f_\Box$ (or $f_\Diamond$), satisfying some conditions,
- every Kripke frame is a BAO,
- modal logic **K** is sound and complete with respect to the class of all BAOs,
- every normal modal logic is sound and complete with respect to some class of BAOs.

Table of Contents

Logic **IL**

Alphabet of logic **IL** is the union of the following sets:

- a countable set $\mathrm{Prop} = \{p_0, p_1, p_2, \dots\}$, of propositional variables,
- a set $\{\bot\}$,
- a set $\{\to\}$,
- a set $\{\rhd\}$ and
- a set $\{(,)\}$.

A formula of **IL** is given by the following:

$$\varphi ::= p \mid \bot \mid \varphi \to \varphi \mid \varphi \rhd \varphi,$$

where $p \in \mathrm{Prop}$.

Other symbols

We define $\neg$, $\wedge$, $\vee$, $\leftrightarrow$, $\top$, $\Box$ i $\Diamond$ as follows:

- $\neg\varphi := \varphi \to \bot$,
- $\varphi \wedge \psi := \neg(\varphi \to \neg\psi)$,
- $\varphi \vee \psi := \neg\varphi \to \psi$,
- $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi)$,
- $\top := \neg\bot$,
- $\Box\varphi := (\neg\varphi) \rhd \bot$ and
- $\Diamond\varphi := \neg\Box\neg\varphi$.

System **IL**

System **IL** contains all propositional tautologies and all instantiations of the following:

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 $\Box(\varphi \to \psi) \to (\varphi \rhd \psi)$,

J2 $((\varphi \rhd \psi) \wedge (\psi \rhd \chi)) \to (\varphi \rhd \chi)$,

J3 $((\varphi \rhd \chi) \wedge (\psi \rhd \chi)) \to ((\varphi \vee \psi) \rhd \chi)$,

J4 $(\varphi \rhd \psi) \to (\Diamond\varphi \to \Diamond\psi)$,

J5 $(\Diamond\varphi \rhd \varphi)$.

Rules of inference are:

- modus ponens: from $\varphi \to \psi$ and φ derive ψ,
- necessitation: from φ derive $\Box\varphi$.

Proof

A *proof* of a formula φ in **IL** is a finite sequence of formulae such that φ is the final formula in the sequence and every formula in the sequence is

- a tautology,
- and instantiation of an axiom schema of **IL**,
- derived by a rule of inference from some of the previous formulas.

If there exists a proof of φ, we refer to φ as *provable* in **IL** or a *theorem* of **IL** and denote it as $\vdash \varphi$.

Derivation

A *derivation* of a formula φ from a set Γ in **IL** is a finite sequence of formulae such that φ is the final formula in the sequence and every formula in the sequence is

- theorem of **IL**,
- an element of Γ,
- derived by modus ponens from some of the previous formulas.

If such a derivation exists, we refer to φ as *derivable* from Γ in **IL** and denote it as $\Gamma \vdash_{\mathbf{IL}} \varphi$.

Table of Contents

Veltman Semantics

Definition

A *Veltman frame* $\mathfrak{F}$ is a triple $(W, R, \{S_w : w \in W\})$, where W is a non-empty set, R transitive and conversely well-founded binary relation on W and $\{S_w : w \in W\}$ a collection of binary relations on $R[w]$, where, for all $w \in W$, S_w is a reflexive and transitive and the restriction of R onto $R[w]$ is contained in S_w.

Definition

A *Veltman model* is a pair $\mathfrak{M} = (\mathfrak{F}, V)$, where $\mathfrak{F}$ is a Veltman frame and $V : \mathrm{Prop} \to \mathcal{P}(W)$ is a *valuation* function.

Veltman Semantics

Valuation indeces a *forcing* relation $\Vdash$, in the following way:

$$
\begin{aligned}
w \Vdash p \quad & \Longleftrightarrow w \in V(p), \\
w \Vdash \bot \quad & \quad\ \text{for no } w \in W, \\
w \Vdash \varphi \to \psi \quad & \Longleftrightarrow w \nVdash \varphi \text{ or } w \Vdash \psi, \\
w \Vdash \Box\varphi \quad & \Longleftrightarrow \forall v(wRv \Rightarrow v \Vdash \varphi), \\
w \Vdash \varphi \rhd \psi \quad & \Longleftrightarrow \forall u\,(wRu\ \&\ u \Vdash \varphi \Rightarrow \exists v(uS_w v\ \&\ v \Vdash \psi))\,.
\end{aligned}
$$

We also write $\mathfrak{M}, w \Vdash \varphi$, if we want to specify the model $\mathfrak{M}$. If for all $w \in W$ in a model $\mathfrak{M}$, $w \Vdash \varphi$ holds, we write $\mathfrak{M} \Vdash \varphi$.
Forcing relation extends the valuation function to a set of all formulae:

$$V(\varphi) = \{w \in W : w \Vdash \varphi\}.$$

Completeness

- F. Veltman, D. de Jongh. *Provability Logics for Relative Interpretability*, Mathematical Logic, Springer, Boston, MA, 1990.
- G. Japaridze, D. de Jongh. *The Logic of Provability*, Handbook of Proof Theory, Elsevier, Amsterdam, 1998.

Theorem (Weak completeness)

If $\nvdash_{\mathbf{IL}} \varphi$, then there exists a finite Veltman model $\mathfrak{M}$ such that $\mathfrak{M} \nVdash \varphi$.

Incompleteness of extensions

E. Goris, J.J. Joosten. *A new principle in the interpretability logic of all reasonable arithmetical theories*, Logic Journal of the IGPL 19, 2011.

Completeness theorem, however, does not hold for extensions of **IL**. Consider systems **IL**P_0 and **IL**R obtained by adding one of the following to our system:

P_0 $\quad (\varphi \rhd \Diamond\psi) \to \Box(\varphi \rhd \psi)$,

R $\quad (\varphi \rhd \psi) \to \neg(\varphi \rhd \neg\chi) \rhd (\psi \rhd \Box\chi)$.

It can be shown that P_0 and R define the same class of frames, but $\nvdash_{\mathbf{IL}\mathsf{P}_0} \mathsf{R}$, which means that system **IL**P_0 is incomplete with respect to Veltman semantics.

Table of Contents

Interpretability algebras

Definition

An *interpretability algebra* is a tuple $\mathfrak{A} = (A, +, -, 0, f_\rhd)$, where $(A, +, -, 0)$ is a boolean algebra and $f_\rhd : A \times A \to A$ is an operator satisfying the following (where we use abbreviations $f_\Box(x) := f_\rhd(-x, 0)$ and $f_\Diamond(x) := -f_\Box(-x)$):

1. $f_\Diamond(0) = 0$,
2. $f_\Box(-f_\Box(x) + x) = f_\Box(x)$,
3. $f_\rhd(x, z) \cdot f_\rhd(y, z) = f_\rhd(x + y, z)$,
4. $-f_\Box(-x + y) + f_\rhd(x, y) = 1$
5. $-f_\rhd(x, y) + (-f_\rhd(y, z)) + f_\rhd(x, z) = 1$,
6. $-f_\rhd(x, y) + f_\Box(-x) + f_\Diamond y = 1$,
7. $f_\rhd(f_\Diamond(x), x) = 1$.

Algebras from frames

Let $\mathfrak{F} = (W, R, \{S_w : w \in W\})$ be a Veltman frame. Then the structure

$$\mathfrak{F}^+ := (\mathcal{P}(W), \cup, ^c, \emptyset, m_\rhd),$$

where

$$m_\rhd(X, Y) = \{w \in W : \forall u \in X\ (wRu \to \exists v \in Y\ (uS_w v))\},$$

is an interpretability algebra, which we refer to as the *complex interpretability algebra* of $\mathfrak{F}$. This means that interpretability algebras are a generalization of Veltman frames.

What about the interpretation of formulae?

Algebraic models

Definition

Let $\mathfrak{A} = (A, +, -, 0, f_{\rhd})$ be an interpretability algebra. An *assignment* is a function $\theta : \mathsf{Prop} \to A$.

We can extend an assignment function to the set of all formulae in the following way:

$$\begin{aligned}
\theta(\bot) &= 0, \\
\theta(\varphi \to \psi) &= -\theta(\varphi) + \theta(\psi), \\
\theta(\varphi \rhd \psi) &= f_{\rhd}(\theta(\varphi), \theta(\psi)).
\end{aligned}$$

When $\mathfrak{A}$ is a complex interpretability algebra of a frame, the assignment is nothing more than a valuation. Hence, interpretability algebras with assignments are a generalization of Veltman models.

Algebraic models

Definition

Let φ and ψ be some formulae. We refer to an expression of the form $\varphi \approx \psi$ as an **IL**-*equation*.
We say that an **IL**-equation $\varphi \approx \psi$ is *true* in an interpretability algebra $\mathfrak{A}$, which we denote as $\mathfrak{A} \vDash \varphi \approx \psi$, if for every assignment θ we have $\theta(\varphi) = \theta(\psi)$.

For complex interpretability algebras, the following holds:

$$\begin{aligned}
(\mathfrak{F}, \theta), w \Vdash \varphi \quad &\text{if and only if} \quad w \in \theta(\varphi), \\
\mathfrak{F} \Vdash \varphi \quad &\text{if and only if} \quad \mathfrak{F}^{+} \vDash \varphi \approx \top, \\
\mathfrak{F} \Vdash \varphi \leftrightarrow \psi \quad &\text{if and only if} \quad \mathfrak{F}^{+} \vDash \varphi \approx \psi.
\end{aligned}$$

Table of Contents

Logic **IL**

Straightforward induction over the length of a derivation may be used to prove the soundness theorem, i.e.

Theorem (Soundness)

If φ is a theorem of **IL**, then $\varphi \approx \top$ is true on any interpretability algebra.

Additionally, we may use the completeness theorem for **IL** with respect to Veltman ssemantics to prove the completeness theorem, i.e.

Theorem (Completeness)

If for every interpretability algebra $\mathfrak{A}$ we have $\mathfrak{A} \vDash \varphi \approx \top$, then $\vdash_{\mathbf{IL}} \varphi$.

Extensions of **IL**

If φ is a formula, we use $\varphi^{\approx}$ to denote the equation $\varphi \approx \top$.

Given a set of formulae Σ, we use V_{Σ} to denote the class of all interpretability algebras in which the set $\{\varphi^{\approx} : \varphi \in \Sigma\}$ is true.

For an extension $\mathbf{IL}^{+}$ of **IL**, we use $V_{\mathbf{IL}^{+}}$ to denote a class of interpretability algebras in which the set of all axioms of $\mathbf{IL}^{+}$ is true.

Extensions of **IL**

We define the relation $\equiv_{\mathbf{IL}^+}$ between formulae in $\mathcal{L}_{\mathbf{IL}}$ in the following way:

$$\varphi \equiv_{\mathbf{IL}^+} \psi \quad \text{if and only if} \quad \vdash_{\mathbf{IL}^+} \varphi \leftrightarrow \psi.$$

If $\varphi \equiv_{\mathbf{IL}^+} \psi$, we say that φ and ψ are equivalent *modulo* $\mathbf{IL}^+$.

Relation $\equiv_{\mathbf{IL}^+}$ is a congruence relation (an equivalence relation which is well behaved with respect to boolean connectives and operators).

Lindenbaum-Tarski $\mathbf{IL}^+$-algebra

Definition

Let $\mathcal{L}_{\mathbf{IL}}$ be the set of all formulae in the language of **IL**. The *Lindenbaum-Tarski* $\mathbf{IL}^+$*-algebra* is the structure

$$\mathfrak{L}_{\mathbf{IL}^+} = (\mathcal{L}_{\mathbf{IL}} / \equiv_{\mathbf{IL}^+}, +, -, 0, f_{\triangleright}),$$

where $[\varphi] + [\psi] := [\varphi \vee \psi]$, $-[\varphi] := [\neg\varphi]$, $0 := [\bot]$ and $f_{\triangleright}([\varphi], [\psi]) := [\varphi \triangleright \psi]$.

Lindenbaum-Tarski $\mathbf{IL}^+$-algebra is an interpretability algebra which belongs to the class $V_{\mathbf{IL}^+}$. Additionally, for all formulae φ,

$$\vdash_{\mathbf{IL}^+} \varphi \quad \text{if and only if} \quad \mathfrak{L}_{\mathbf{IL}^+} \models \varphi \approx \top.$$

Completeness

Lindenbaum-Tarski $\mathbf{IL}^+$-algebra falsifies all non-theorems of $\mathbf{IL}^+$. Therefore, if we have a non-theorem of $\mathbf{IL}^+$, we have found an algebra belonging to the class $V_{\mathbf{IL}^+}$ which falsifies it. Hence, we have the following theorem:

Theorem (Completeness for extensions)

Let $\mathbf{IL}^+$ be an extension of $\mathbf{IL}$. Then $\mathbf{IL}^+$ is sound and complete with respect to $V_{\mathbf{IL}^+}$, i.e.,

$$\vdash_{\mathbf{IL}^+} \varphi \quad \text{if and only if} \quad V_{\mathbf{IL}^+} \models \varphi^{\approx}.$$

Table of Contents

Properties of interpretability algebras

- a boolean algebra together with an operator $f_{\rhd}$, satisfying some conditions,
- every Veltman frame is an interpretability algebra,
- interpretabiltiy logic **IL** is sound and complete with respect to the class of all interpretability algebras,
- every extension of **IL** is sound and complete with respect to some class of interpretability algebras.

Concluding remarks

This result

- improves upon Veltman semantics, in a natural way analogous to modal logic,
- proves completeness of extensions in a modular way,
- generalizes the notion of BAO to interpretability logics.

Concluding remarks

Future work:

- general frame semantics,
- duality theory.

Concluding remarks

Thank you for your attention!

Bibliography

- Blackburn P., de Rijke, M., Venema, Y.: Modal Logic, Cambridge University Press (2001)
- Goris, E., Joosten, J.J.: A new principle in the interpretability logic of all reasonable arithmetical theories, Logic Journal of the IGPL 19, 1–17 (2011)
- Goris E., Joosten J.J.: Modal matters in interpretability logics, Logic Journal of the IGPL 16, 371–412 (2008)
- Šestak, T.: Algebraic semantics for interpretability logics. In: V. Gradimondo, E.E. Rosina (eds.): Preliminary Proceedings of the ESSLLI 2025 Student Session, pp. 2–11. (2025)
- Veltman, F., de Jongh, D.: Provability logics for relative interpretability. In: P.P. Petkov (ed.), Mathematical Logic, Proceedings of the 1988 Heyting Conference, pp. 31–42. Plenum Press, (1990)

Bibliography

- Visser, A.: An overview of interpretability logic. In: M. Kracht et al. (eds.), Advances in Modal Logic, Vol. 1, pp. 307–359. CSLI Publications (1998)
- Visser, A.: Interpretability logic. In: P.P. Petkov (ed.), Mathematical Logic, Proceedings of the 1988 Heyting Conference, pp. 175–210. Plenum Press (1990)
- Visser, A., Preliminary Notes on Interpretability Logic, Logic Group Preprint Series nr. 29, Department of Philosophy, University of Utrecht (1988)