The Preservativity Logic

S. Mojtaba Mojtahedi

University of Tehran

Joint work with: Mohammad Ardeshir

May 12, 2016

The languages

Propositional languages: $\mathcal{L}_0$, $\mathcal{L}_\Box$ and $\mathcal{L}_\rhd$.

- $\mathcal{L}_0 := \{\wedge, \vee, \rightarrow, \bot\}$ and $\neg A$ is defined as $A \rightarrow \bot$.
- $\mathcal{L}_\Box := \mathcal{L}_0$ with the unary modal operator $\Box$.
- $\mathcal{L}_\rhd := \mathcal{L}_0$ with the addition of the binary modal operator $\rhd$.

The theories IPC and CPC are intuitionistic and classical propositional calculi.

First-order language of arithmetic: $S, +, ., =, 0$

PA and HA

The theory PA is the first-order axiomatization of number theory and HA is the same theory over intuitionistic logic instead of classical logic.

Provability Logic

Roughly speaking, the Provability Logic of a first-order theory T, $\mathcal{PL}(T)$, is the set of all *valid* (in T) modal propositions, in which $\Box$ interpreted as provability predicate. T should be strong enough to define provability predicate $\mathrm{Pr}_T(x)$.

More precisely, $\mathcal{PL}(T)$ contains all $A \in \mathcal{L}_\Box$ such that for all arithmetical interpretations $(.)^*$, $T \vdash A^*$. $(.)^*$ commutes with all connectives except for $\Box$:

$$(\Box A)^* := \mathrm{Pr}_T(\ulcorner A^* \urcorner)$$

Why it is interesting?

From a philosophical point of view, provability logic is interesting because:

- The concept of provability in a fixed theory of arithmetic has a unique and non-problematic meaning, other than concepts like necessity and knowledge studied in modal and epistemic logic.
- Provability logic provides tools to study the notion of self-reference.

Review of results

- Gödel 1931: $\neg\Box\bot \notin \mathcal{PL}(\mathsf{PA})$
- Löb 1953: $\mathsf{Lob} := \Box(\Box A \to A) \to \Box A \in \mathcal{PL}(\mathsf{PA})$
- Leivant 1975: $\Box(A \vee B) \to \Box(\Box A \vee B) \in \mathcal{PL}(\mathsf{HA})$
- Solovay 1976: $\mathcal{PL}(\mathsf{PA}) = \mathsf{K4} + \mathsf{Lob} = \mathsf{GL}$
- Visser 1981: $\Box\neg\neg\Box A \to \Box\Box A \in \mathcal{PL}(\mathsf{HA})$
- Visser 1981: Provides decision algorithm for closed fragment (i.e. modal propositions with no atomic variables) of $\mathcal{PL}(\mathsf{HA})$
- Iemhoff 2001: Provides some uniform axiomatization in an extended language and Kripke model-theory for all the known principles of $\mathcal{PL}(\mathsf{HA})$.
- We do not know that $\mathcal{PL}(\mathsf{HA})$ is decidable.

The Interpretations of first-order theories

Interpretations

We say that i is an interpretation of S in T (or T interprets S via i or $T \rhd_i S$) if for all φ with $S \vdash \varphi$ we have $T \vdash \varphi^i$.

Example

$\mathsf{ZFC} \rhd_{int} \mathsf{PA}$ and $\mathsf{PA} \rhd_{int} I\Sigma_n$

Formalization of interpretability (Orey)

$$T \rhd_{int} S \Leftrightarrow \forall x \Box_T \mathsf{Con}(S_x)$$

The Interpretability Logic

The Interpretability logic of PA

$A \rhd_{int} B$ iff $\mathsf{PA} + A$ interprets the theory $\mathsf{PA} + B$.

The Theory ILM

- $\mathsf{GL} := \mathsf{K4} + \Box(\Box A \to A) \to \Box A$, $(\Box A := \neg A \rhd \bot)$
- $\Box(A \to B) \to (A \rhd B)$,
- $A \rhd B \wedge B \rhd C \to A \rhd C$,
- $[(B \rhd A) \wedge (C \rhd A)] \to [(B \vee C) \rhd A]$,
- $A \rhd B \to (\Diamond A \to \Diamond B)$,
- $\Diamond A \rhd A$,
- $A \rhd B \to (A \wedge \Box C) \rhd (B \wedge \Box C)$. (Montagna's Principle)

The Interpretability Logic of PA

The interpretability logic of PA:=

$$\mathcal{IL}(\mathsf{PA}) := \{A \in \mathcal{L}_\rhd | \forall * (\mathsf{PA} \vdash A^*)\}$$

in which $(A \rhd B)^* := \forall x \Box_{\mathsf{PA}} (A^* \to \mathsf{Con}(\mathsf{PA}_x + B^*))$

Alessandro Berarducci- Vladimir Shavrukov 1990

$\mathsf{ILM} = \mathcal{IL}(\mathsf{PA})$.

Σ_1-Formulas and Π_1-formulas

Σ_1-formulas are exactly those formulas $A(x)$ such that the set $\{x|A(x)\}$ is recursively enumerable.

More precisely $A(x) = \exists y B(x, y)$, such that every quantifier in B is bounded, i.e. in the form $\exists x \leq t$ or $\forall x \leq t$.
A Π_1-formula is the negation of a Σ_1-formula.

The Conservativity Logic

- $A \rhd_{conser} B \equiv \forall C \in \Pi_1(\Box(A \to C) \to \Box(B \to C))$.
- $\mathcal{CL}(\mathsf{PA}) := \{A \in \mathcal{L}_{\rhd} | \forall * \mathsf{PA} \vdash A^*\}$.
- In above item $(A \rhd B)^* := \forall C \in \Pi_1(\Box(A^* \to C^*) \to \Box(B^* \to C^*))$ and $*$ commutes with other connectives.
- $\Box A := \bot \rhd \neg A$

Petr Hajék and Franco Montagna 1990

$\mathcal{CL}(\mathsf{PA}) = \mathcal{IL}(\mathsf{PA}) = ILM$.

The Σ-Preservativity Logic

$$A \rhd_{pre} B \equiv \forall C \in \Sigma_1(\Box(C \to A) \to \Box(C \to B))$$

In classical theories we have

$$A \rhd_{pre} B \quad \text{iff} \quad \neg A \rhd_{conser} \neg B$$

The preservativity logic:= $\mathcal{PSL}(T) := \{A \in \mathcal{L}_\rhd | \forall * T \vdash A^*\}$

The Preservativity Logic of HA : Known Axioms iPH

- $\Box A \equiv \top \rhd A$,

Taut All Tautologies of IPC ,

P1 $(A \rhd B \wedge B \rhd C) \to A \rhd C$,

P2 $(A \rhd B \wedge A \rhd C) \to A \rhd (B \wedge C)$,

Dp $(B \rhd A \wedge C \rhd A) \to (B \vee C) \rhd A$,

4p $A \rhd \Box A$,

Lp $(\Box A \to A) \rhd A$,

Mp $A \rhd B \to (C \to A) \rhd (C \to B)$, For Boxed propositions $C := \Box C'$,

Vp $(A \to (F_{n+1} \vee F_{n+2})) \rhd \{B\}(F_1, \dots, F_{n+2})$ in which

$$A = \bigwedge_{i=1}^{n} (E_i \to F_i) \quad \{B\}(C_1, \dots, C_n) := \bigvee_{i=1}^{n} \{B\}(C_i)$$

Some results

- Visser and de Jongh showed that iPH is sound for arithmetical interpretations in HA . (1994)
- Iemhoff proved completeness of iPH for some frame property of Kripke models. (PhD Thesis, 2001)
- Iemhoff Conjectured that iPH is the preservativity logic of HA, i.e. iPH is also complete for arithmetical interpretations in HA.
- Iemhoff showed that based on iPH one could characterize the admissible rules of IPC. (PhD Thesis, 2001)
- Visser showed that based on iPH one could characterize the propositional admissible rules of HA. (2002)

Preservativity Logic of HA for Σ-substitutions: iPH^σ

- $\Box A \equiv \top \rhd A$,

Taut All Tautologies of IPC , $p \to \Box p$

P1 $(A \rhd B \wedge B \rhd C) \to A \rhd C$,

P2 $(A \rhd B \wedge A \rhd C) \to A \rhd (B \wedge C)$,

Dp $(B \rhd A \wedge C \rhd A) \to (B \vee C) \rhd A$,

4p $A \rhd \Box A$,

Lp $(\Box A \to A) \rhd A$,

Mp $A \rhd B \to (C \to A) \rhd (C \to B)$, For Boxed propositions $C := \Box C'$ and also for atomic C,

Vp $(A \to (F_{n+1} \vee F_{n+2})) \rhd [B](F_1, \ldots, F_{n+2})$ in which

$$A = \bigwedge_{i=1}^{n} (E_i \to F_i) \quad \{B\}(C_1, \ldots, C_n) := \bigvee_{i=1}^{n} [B](C_i)$$

Results on iPH^σ

Visser 2002

The propositional part of the admissible rules of HA for Σ-substitutions are all $A \rhd B \in \mathsf{iPH}^\sigma$ in which $A, B \in \mathcal{L}_0$.

Ardeshir-Mojtahedi 2013

The provability logic of HA for Σ-substitutions is

$$\mathsf{iH}^\sigma := \{A \in \mathcal{L}_\Box | \mathsf{iPH}^\sigma \vdash A\}$$

Moreover $\mathsf{iPH}^\sigma \vdash A$ is decidable for $A \in \mathcal{L}_\Box$.

TNNIL

We call a formula in $\mathcal{L}_\Box$ to be TNNIL , if there is no $\to$ in the left side of a $\to$, except in the scope of some $\Box$. For example the following propositions are TNNIL:

- $\Box(p \to \bot) \to q$
- $\Box\bot \to (p \to q)$

And the following are not:

- $(p \to \bot) \to \bot$
- $\Box((p \to \bot) \to \bot) \to q$

Visser's Algorithm

Visser in his PhD thesis invented an algorithm such that assign a TNNIL proposition to each $\mathcal{L}_0$-proposition, such that it is best TNNIL approximation from below (in IPC). We can extend this method to the language $\mathcal{L}_\Box$ in an straightforward way. Let A^+ denote the result of TNNIL algorithm for A as input. Let's see some examples:

- $(\neg\neg p)^+ = p$. As a result of Visser's Theorem, this means that $\mathsf{IPC} \vdash p \to \neg\neg p$ and moreover for any TNNIL proposition $A \in \mathcal{L}_0$ such that $\mathsf{IPC} \vdash A \to \neg\neg p$, we have $\mathsf{IPC} \vdash A \to p$.
- $(\neg p \to q)^+ = p \vee q$
- $(\neg\neg\Box(\neg p \to q))^+ = \Box(p \vee q)$

The theory LC

LC has the following axioms and rules:

- Axioms and rules of $i\mathsf{K4} := \mathsf{IPC} + \Box(A \to B) \to (\Box A \to \Box B) + \Box A \to \Box\Box A.$
- Completeness: $A \to \Box A.$
- Löb's axiom: $\Box(\Box A \to A) \to \Box A.$

The theory LC

LC has the following axioms and rules:

- Axioms and rules of
  $i\mathsf{K4} := \mathsf{IPC} + \Box(A \to B) \to (\Box A \to \Box B) + \Box A \to \Box\Box A.$
- Completeness: $A \to \Box A$.
- Löb's axiom: $\Box(\Box A \to A) \to \Box A$.

Finite Model Property for LC

LC is sound and complete for finite *perfect* Kripke models (definition comes next) and hence it is decidable.

The theory LC

LC has the following axioms and rules:

- Axioms and rules of $i\mathsf{K4} := \mathsf{IPC} + \Box(A \to B) \to (\Box A \to \Box B) + \Box A \to \Box\Box A$.
- Completeness: $A \to \Box A$.
- Löb's axiom: $\Box(\Box A \to A) \to \Box A$.

Finite Model Property for LC

LC is sound and complete for finite *perfect* Kripke models (definition comes next) and hence it is decidable.

A salient fact

LC and iPH^σ prove the same TNNIL-propositions.

Decision Algorithm for $\mathcal{PL}_{\Sigma}(\mathsf{HA})$

Let A be given in a modal language. We decide $A \in \mathcal{PL}_{\Sigma}(\mathsf{HA})$ in the following 2 steps:

1. Compute A^{+}, by Visser's algorithm.
2. Decide $\mathsf{LC} \vdash A^{+}$. If $\mathsf{LC} \vdash A^{+}$ then we have $A \in \mathcal{PL}_{\Sigma}(\mathsf{HA})$, else $A \notin \mathcal{PL}_{\Sigma}(\mathsf{HA})$.

Main Theorem

Let $\mathsf{LC} \nvdash A$ for some TNNIL proposition $A \in \mathcal{L}_{\Box}$. Then there exists some Σ_1-substitution such that $\mathsf{HA} \nvdash A^{*}$.

Perfect Kripke models

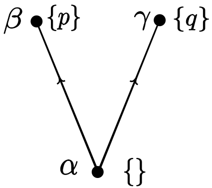
Let $\mathcal{K} = (K, \leq, \mathcal{R}, \Vdash)$. We call $\mathcal{K}$ to be a *perfect* Kripke model if

- $(K, \leq)$ is a partially ordered set,
- $\mathcal{R}$ is an irreflexive binary relation over K, (Löb's axiom)
- $\mathcal{R} \subseteq \leq$, (Completeness axiom)
- $(\leq; \mathcal{R}) \subseteq \mathcal{R}$, in which $u(\leq; \mathcal{R})v$ iff $u \leq w \ \mathcal{R} \ v$ for some $w \in K$, (Guaranties the persistence of $A \rhd B$ over Kripke model)
- $u \Vdash p$ and $u \leq v$ implies $v \Vdash p$ for atomic p.

We extend $\Vdash$ to the modal language in the following way:

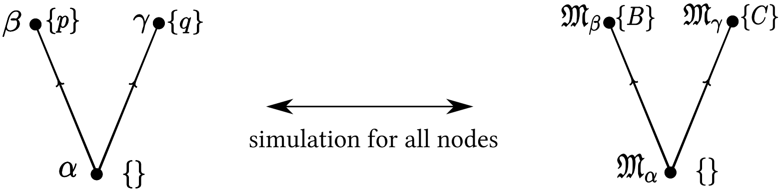
- $u \Vdash A \to B$ iff for all $v \geq u$, $v \Vdash A$ implies $v \Vdash B$,
- $u \Vdash \Box A$ iff for all v such that $u \ \mathcal{R} \ v$, we have $v \Vdash A$.

$$A := (p \to q) \vee (q \to p)$$



$$\beta \Vdash p \quad , \quad \beta \nVdash q \quad , \quad \gamma \Vdash q \quad , \quad \gamma \nVdash p$$

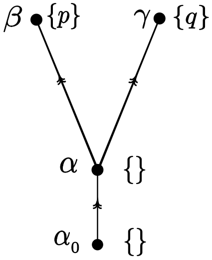
$$\alpha \leq \beta, \gamma \quad , \quad \alpha \nVdash p, q$$



$$B := \exists x(F(x) = \beta) \quad \text{and} \quad C := \exists x(F(x) = \gamma)$$

$$\mathfrak{M}_\delta \models T_\delta \quad , \quad T_\delta := \mathsf{PA} + (\lim_{x\to\infty} F(x) = \delta)$$

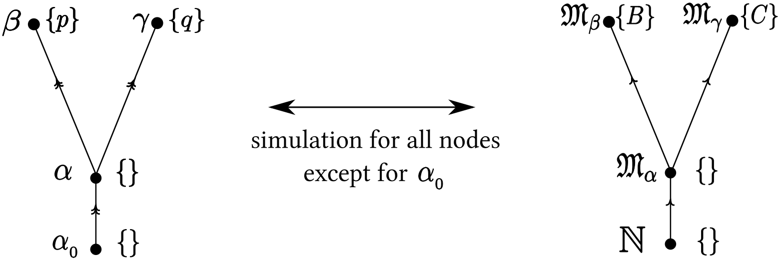
$$A = \Box(p \vee q) \to (\Box p \vee \Box q)$$



$\beta \Vdash p,\ \beta \nVdash q,\ \gamma \Vdash q,\ \gamma \nVdash p$

$\alpha\, \mathcal{R}\, \beta, \gamma\ , \quad \alpha \leq \beta, \gamma$

$\alpha_0\, \mathcal{R}\, \alpha, \beta, \gamma, \quad \alpha_0 \leq \alpha, \beta, \gamma, \quad \alpha, \alpha_0 \nVdash p, q$

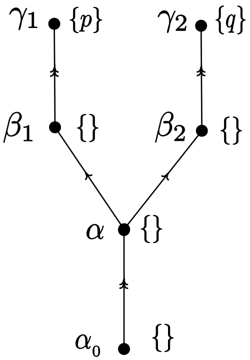


$$B := \exists x(F(x) = \beta) \quad \text{and} \quad C := \exists x(F(x) = \gamma)$$

$$T_\delta := \mathsf{PA} + (\lim_{x\to\infty} F(x) = \delta) + \mathsf{Prov}_{\mathsf{HA}}(\ulcorner \varphi_\delta \urcorner)$$

$$\varphi_\alpha := B \vee C, \varphi_\beta := \varphi_\gamma := B \wedge C$$

$$A := \Box(p \vee q) \to [(\Box p \to (p \vee q \vee \Box q)) \vee (\Box q \to (p \vee q \vee \Box p))]$$



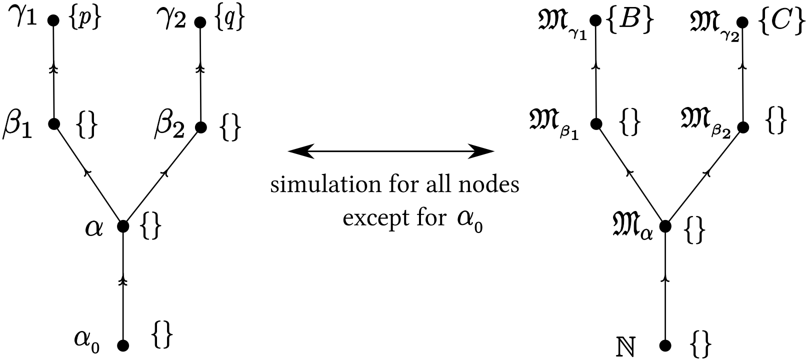
$\gamma_1 \Vdash p$ $\gamma_2 \Vdash q$

$\alpha_0, \alpha, \beta_1, \beta_2, \gamma_2 \nVdash p$ $\quad \alpha_0, \alpha, \beta_1, \beta_2, \gamma_1 \nVdash q$

$\alpha_0 \leq \alpha, \beta_1, \beta_2, \gamma_1, \gamma_2$ $\quad \alpha_0 \,\mathcal{R}\, \alpha, \beta_1, \beta_2, \gamma_1, \gamma_2$

$\alpha \leq \beta_1, \beta_2, \gamma_1, \gamma_2$ $\quad \alpha \,\mathcal{R}\, \gamma_1, \gamma_2$ $\quad \beta_i \leq \gamma_i$

$\beta_i \,\mathcal{R}\, \gamma_i$ $\quad \alpha \,\not\mathcal{R}\, \beta_1$ $\quad \alpha \,\not\mathcal{R}\, \beta_2$



$$B := (\exists x F(x) = \gamma_1) \quad C := (\exists x F(x) = \gamma_2)$$

$$\varphi_{\alpha} := B \vee C \quad \varphi_{\beta_1} := B \quad \varphi_{\beta_2} := C$$

$$\varphi_{\gamma_1} := \varphi_{\gamma_2} := B \wedge C \quad T_\delta := \mathsf{PA} + \lim_{x \to \infty} F(x) = \delta + \mathsf{Prov}_{\mathsf{HA}}(\ulcorner \varphi_\delta \urcorner)$$

$F(0) := \alpha_0$. Let $F(n) := \delta$, define $F(n+1) := \delta'$ if one of the following cases occurs, otherwise we define $F(n+1) := F(n) = \delta$.

- $\delta \,\mathcal{R}\, \delta'$ and there exists some witness (which is less than or equal to $n+1$) for the inconsistency of $T_{\delta'}$, or in other words, there exists some proof (in PA) with the Gödel number $\leq n+1$ for the statement

$$\neg[\lim_{x\to\infty} F(x) = \delta' \wedge \mathsf{Prov}_{\mathsf{HA}}(\ulcorner \varphi_{\delta'} \urcorner)]$$

- All of the following conditions hold:
  - $\delta \not\mathcal{R} \delta'$ and $\delta \leq \delta'$,
  - There exists some witness (which is less than or equal to $n+1$) for the inconsistency of $T_{\delta'}$,
  - The $n+1$-inconsistency rank of $T_{\delta'}$ (we call it $r(\delta', n+1)$) is less than the $n+1$-inconsistency rank of T_δ (we call it $r(\delta, n+1)$),
  - $F(r(\delta', n+1)) \,\mathcal{R}\, \delta$.

The Inconsistency rank

The inconsistency rank of T_δ is defined to be the minimum k such that there exists a witness (less than or equal to $n+1$) for the inconsistency of

$$\mathsf{PA}_k + \lim_{x\to\infty} F(x) = \delta + \mathsf{Prov}_{\mathsf{HA}}(\ulcorner \varphi_\delta \urcorner)$$

In above definition, PA_k is the theory $I\Sigma_1$ plus induction axiom for those formulas with Gödel number less than k.

Thanks For Your Attention