The BHK Interpretation: Looking through Gödel's Classical Lens

Amir Akbar Tabatabai

Faculty of Humanities, Utrecht University

LC 2019, Prague

A Very Old Problem

Consider the modal system **S4**:

- Axiom **K**: $\Box(A \to B) \to (\Box A \to \Box B)$,
- Axiom **T**: $\Box A \to A$,
- Axiom **4**: $\Box A \to \Box\Box A$,

A Very Old Problem

Consider the modal system **S4**:

- Axiom **K**: $\Box(A \to B) \to (\Box A \to \Box B)$,
- Axiom **T**: $\Box A \to A$,
- Axiom **4**: $\Box A \to \Box\Box A$,

– **MP**: If A and $A \to B$ are provable then B is also provable,

– **NC**: If A is provable then $\Box A$ is also provable.

A Very Old Problem

Consider the modal system **S4**:

- Axiom **K**: $\Box(A \to B) \to (\Box A \to \Box B)$,
- Axiom **T**: $\Box A \to A$,
- Axiom **4**: $\Box A \to \Box\Box A$,

– **MP**: If A and $A \to B$ are provable then B is also provable,

– **NC**: If A is provable then $\Box A$ is also provable.

If we interpret $\Box$ as *informal provability* then all the axioms and rules are valid.

A Very Old Problem

Consider the modal system **S4**:

- Axiom **K**: $\Box(A \to B) \to (\Box A \to \Box B)$,
- Axiom **T**: $\Box A \to A$,
- Axiom **4**: $\Box A \to \Box\Box A$,

– **MP**: If A and $A \to B$ are provable then B is also provable,

– **NC**: If A is provable then $\Box A$ is also provable.

If we interpret $\Box$ as *informal provability* then all the axioms and rules are valid.

Gödel's 1933 Problem

Is it possible to formalize this *informal provability interpretation* using some *concrete* classical proofs?

The Simplest Approach

The most natural attempt is fixing a strong enough r.e. theory T and interpret $\Box$ as a natural provability predicate for the theory T.

This interpretation is not sound because by Necessitation and the axiom **T**, we have $\mathbf{S4} \vdash \Box(\Box\bot \to \bot)$ while its interpretation will be $\mathrm{Pr}_T(\neg \mathrm{Pr}_T(\bot))$. But T can not prove its own consistency.

The Simplest Approach

The most natural attempt is fixing a strong enough r.e. theory T and interpret $\Box$ as a natural provability predicate for the theory T.

This interpretation is not sound because by Necessitation and the axiom **T**, we have $\mathbf{S4} \vdash \Box(\Box\bot \to \bot)$ while its interpretation will be $\mathrm{Pr}_T(\neg \mathrm{Pr}_T(\bot))$. But T can not prove its own consistency.

Where is the clash between the previous interpretation and the intuitive interpretation?

The Simplest Approach

The most natural attempt is fixing a strong enough r.e. theory T and interpret $\Box$ as a natural provability predicate for the theory T.

This interpretation is not sound because by Necessitation and the axiom **T**, we have $\mathbf{S4} \vdash \Box(\Box\bot \to \bot)$ while its interpretation will be $\mathrm{Pr}_T(\neg \mathrm{Pr}_T(\bot))$. But T can not prove its own consistency.

Where is the clash between the previous interpretation and the intuitive interpretation?

In the formula $\Box(\Box\bot \to \bot)$, the inner box refers to the provability in a theory T, but the outer box refers to the provability in the meta-theory of T which is not necessarily equal to T itself.

A More Sophisticated Approach

In this sense the natural interpretation of modal formulas needs:

- a model M capturing the real world and,
- a hierarchy of theories $\{T_n\}_{n=0}^{\infty}$ capturing the whole hierarchy of theories, meta-theories, meta-meta-theories and so on.

A More Sophisticated Approach

In this sense the natural interpretation of modal formulas needs:

- a model M capturing the real world and,
- a hierarchy of theories $\{T_n\}_{n=0}^{\infty}$ capturing the whole hierarchy of theories, meta-theories, meta-meta-theories and so on.

Definition

A provability model is a pair $(M, \{T_n\}_{n=0}^{\infty})$ where M is a model of $I\Sigma_1$ and $\{T_n\}_{n=0}^{\infty}$ is a hierarchy of arithmetical r.e. theories such that for any n, $I\Sigma_1 \subseteq T_n \subseteq T_{n+1}$ provably in $I\Sigma_1$.

A More Sophisticated Approach

In this sense the natural interpretation of modal formulas needs:

- a model M capturing the real world and,
- a hierarchy of theories $\{T_n\}_{n=0}^{\infty}$ capturing the whole hierarchy of theories, meta-theories, meta-meta-theories and so on.

Definition

A provability model is a pair $(M, \{T_n\}_{n=0}^{\infty})$ where M is a model of $I\Sigma_1$ and $\{T_n\}_{n=0}^{\infty}$ is a hierarchy of arithmetical r.e. theories such that for any n, $I\Sigma_1 \subseteq T_n \subseteq T_{n+1}$ provably in $I\Sigma_1$.

A provability model $(M, \{T_n\}_{n=0}^{\infty})$ is called reflexive if for any n, M thinks that T_n is sound and $T_{n+1} \vdash \mathrm{Rfn}(T_n)$. We will denote this class by **Ref**.

Definition

By a witness w for a formula A, we mean a sequence that assigns numbers to occurrences of the boxes in the formula A such that the number for an outer box is greater than all the numbers assigned to the inner boxes.

Definition

By a witness w for a formula A, we mean a sequence that assigns numbers to occurrences of the boxes in the formula A such that the number for an outer box is greater than all the numbers assigned to the inner boxes.

Example

For instance, $(2, 1)$ is a witness for $\Box(\Box p \to p)$ while $(0, 1)$ and $(3, 3)$ are not.

Definition

By a witness w for a formula A, we mean a sequence that assigns numbers to occurrences of the boxes in the formula A such that the number for an outer box is greater than all the numbers assigned to the inner boxes.

Example

For instance, $(2,1)$ is a witness for $\Box(\Box p \to p)$ while $(0,1)$ and $(3,3)$ are not.

Definition

Let w be a witness for A and σ an arithmetical substitution which assigns an arithmetical sentence to a propositional variable. And also let $(M, \{T_n\}_{n=0}^{\infty})$ be a provability model. By $A^\sigma(w)$ we mean an arithmetical sentence which results by substituting the variables by the values of σ and interpreting any box as the provability predicate of T_n if the corresponding number in the witness for this box was n. The interpretation of boolean connectives are themselves.

Example

Let $(M, \{T_n\}_{n=0}^{\infty})$ be a reflexive provability model. Then the formula $A = \Box(\Box p \to p)$ is true in this model. It is enough to pick the witness $w = (1, 0)$. Then the interpretation of the formula under the arithmetical interpretation σ is $A^\sigma(w) = \mathrm{Pr}_{T_1}(\mathrm{Pr}_{T_0}(p^\sigma) \to p^\sigma)$ which is true in M.

Example

Let $(M, \{T_n\}_{n=0}^{\infty})$ be a reflexive provability model. Then the formula $A = \Box(\Box p \to p)$ is true in this model. It is enough to pick the witness $w = (1, 0)$. Then the interpretation of the formula under the arithmetical interpretation σ is $A^\sigma(w) = \mathrm{Pr}_{T_1}(\mathrm{Pr}_{T_0}(p^\sigma) \to p^\sigma)$ which is true in M.

A Conjectured Soundness-Completeness Theorem

S4 $\vdash A$ iff there exists a witness for A such that all arithmetical interpretations of A in all reflexive models hold, i.e.,

$$\mathbf{S4} \vdash A \Longleftrightarrow \exists w \forall \sigma \forall (M, \{T_n\}_{n=0}^{\infty}) \in \mathbf{Ref}\ M \vDash A^\sigma(w).$$

Example

Let $(M, \{T_n\}_{n=0}^{\infty})$ be a reflexive provability model. Then the formula $A = \Box(\Box p \to p)$ is true in this model. It is enough to pick the witness $w = (1, 0)$. Then the interpretation of the formula under the arithmetical interpretation σ is $A^\sigma(w) = \mathrm{Pr}_{T_1}(\mathrm{Pr}_{T_0}(p^\sigma) \to p^\sigma)$ which is true in M.

A Conjectured Soundness-Completeness Theorem

S4 $\vdash A$ iff there exists a witness for A such that all arithmetical interpretations of A in all reflexive models hold, i.e.,

$$\mathbf{S4} \vdash A \iff \exists w \forall \sigma \forall (M, \{T_n\}_{n=0}^{\infty}) \in \mathbf{Ref}\ M \vDash A^\sigma(w).$$

The $\exists w$ is based on the assumption that there were valid indices by which we informally argued but now we have forgotten them.

The Herbrand Phenomenon

Unfortunately, this conjecture does not hold. For instance while the formula $\neg\Box(\neg\Box p \wedge p)$ is provable in **S4**, it has no witness that works for all reflexive provability models.

The Herbrand Phenomenon

Unfortunately, this conjecture does not hold. For instance while the formula $\neg\Box(\neg\Box p \wedge p)$ is provable in **S4**, it has no witness that works for all reflexive provability models.

The reason is different roles that on box can play. Our interpretation assumes there was only one index for any box that we have forgotten and we want to remember. This is not true. Think about the formula $\neg\Box_2(\neg\Box_1 p \wedge p) \vee \neg\Box_1(\neg\Box_0 p \wedge p)$. If we forget the indices, then we have $\neg\Box(\neg\Box p \wedge p) \vee \neg\Box(\neg\Box p \wedge p)$ which is equivalent to $\neg\Box(\neg\Box p \wedge p)$. But based on our interpretation, when we want to remember the index, it can be $\neg\Box_2(\neg\Box_1 p \wedge p)$ or $\neg\Box_1(\neg\Box_0 p \wedge p)$, and not their disjunction.

Expansions

To capture these different roles we introduce expansions. They are similar to expansions in the generalized Herbrand's theorem.

Definition

$E(A)$, the set of all expansions of A, is inductively defined as follows:

- If A is an atom, $E(A) = \{A\}$.
- If $A = B \circ C$, then $E(A) = \{D \circ E \mid D \in E(B) \text{ and } E \in E(C)\}$ for $\circ \in \{\wedge, \vee, \rightarrow\}$.
- If $A = \neg B$, then $E(A) = \{\neg D \mid D \in E(B)\}$.
- If $A = \Box B$, then $E(A) = \{\Box \bigvee_{i=1}^{k} D_i \mid \forall 1 \leq i \leq k,\ D_i \in E(B)\}$.

Informally speaking, an expansion of a formula A is a formula resulted by replacing any formula after a box with disjunctions of the expansions of the formula. For instance, $\Box(\Box p \vee \Box p)$ is an expansion for $\Box\Box p$.

The Main Theorem

Soundness-Completeness Theorem

S4 $\vdash A$ iff there exist finite number of expansions of A like $B_1, \ldots, B_k$, a witness for $\bigvee_{i=1}^{k} B_i$ such that all arithmetical interpretations of $\bigvee_{i=1}^{k} B_i$ in all reflexive models hold, i.e.,

$$\mathbf{S4} \vdash A \Longleftrightarrow \exists w \exists B_1, \ldots B_k \forall \sigma \forall (M, \{T_n\}_{n=0}^{\infty}) \in \mathbf{Ref}\ M \vDash (\bigvee_{i=1}^{k} B_i)^{\sigma}(w).$$

The Main Theorem

Soundness-Completeness Theorem

S4 $\vdash A$ iff there exist finite number of expansions of A like $B_1, \ldots, B_k$, a witness for $\bigvee_{i=1}^{k} B_i$ such that all arithmetical interpretations of $\bigvee_{i=1}^{k} B_i$ in all reflexive models hold, i.e.,

$$\mathbf{S4} \vdash A \iff \exists w \exists B_1, \ldots B_k \forall \sigma \forall (M, \{T_n\}_{n=0}^{\infty}) \in \mathbf{Ref}\ M \vDash (\bigvee_{i=1}^{k} B_i)^{\sigma}(w).$$

Proof.

For soundness, use the cut-free system for **S4**. For completeness, use a modification of Solovay's technique. □

The BHK Interpretation and Gödel's Classical Lens

- BHK interpretation interprets the connectives as operations on some informal open notion of *Proof*.
- What are these proofs? Gödel proposed using *classical* proofs.

The BHK Interpretation and Gödel's Classical Lens

- BHK interpretation interprets the connectives as operations on some informal open notion of *Proof*.
- What are these proofs? Gödel proposed using *classical* proofs.
- He reinvent the system **S4** as a calculus for classical provability and using BHK as a base for his translation:
  - $p^b = \Box p$ and $\bot^b = \Box\bot$
  - $(A \circ B)^b = A^b \circ B^b$ for $\circ \in \{\wedge, \vee\}$
  - $(A \to B)^b = \Box(A^b \to B^b)$

  claimed that this interpretation is sound and complete for **IPC**, i.e, **IPC** $\vdash A$ iff **S4** $\vdash A^b$.

A Formalization for BHK Interpretation via Classical Proofs

One problem remains open. What is the concrete provability interpretation of **S4** based on concrete proofs?

A Formalization for BHK Interpretation via Classical Proofs

One problem remains open. What is the concrete provability interpretation of **S4** based on concrete proofs? Combining our provability interpretation with Gödel's translation, we will have a formalization for the BHK interpretation:

Soundness-Completeness Theorem

IPC $\vdash A$ iff there exist finite number of expansions of A^b like $B_1, \ldots, B_k$, a witness for $\bigvee_{i=1}^k B_i$ such that all arithmetical interpretations of $\bigvee_{i=1}^k B_i$ in all reflexive models hold, i.e.,

$$\mathbf{IPC} \vdash A \iff \exists w \exists B_1, \ldots B_k \forall \sigma \forall (M, \{T_n\}_{n=0}^{\infty}) \in \mathbf{Ref}\ M \vDash (\bigvee_{i=1}^{k} B_i)^{\sigma}(w).$$

More Characterizations

Modal	Propositional	Provability Models
K4	**BPC**	All Models
KD4	**EBPC**	Consistent Models
S4	**IPC**	Reflexive Models
GL	**FPL**	Constant Models
Above **KD45**	**CPC**	No Models

More Characterizations

Modal	Propositional	Provability Models
K4	**BPC**	All Models
KD4	**EBPC**	Consistent Models
S4	**IPC**	Reflexive Models
GL	**FPL**	Constant Models
Above **KD45**	**CPC**	No Models

- Using hierarchies provides a framework to generalize Solovay's result to capture different modal logics.
- Since in all the propositional results the Gödel's translation (the BHK interpretation) is fixed the result suggests that believing only in BHK interpretation, there could be different equally valid *intuitionistic logics* rather than *the* intuitionitic logic. The difference between these logics is in the ontological commitments that we put on our meta-theories.

Thank you for your attention!