# A Carnapian approach to the meaning of logical constants: the case of modal logic

Dag Westerståhl
Stockholm University

(joint work with Denis Bonnay, Université Paris Ouest)

Berkeley Logic Colloquium
UC Berkeley, October 7, 2016

# Logicality

What makes a word/symbol logical?

- model-theoretic approach: invariance
  - iso-invariance necessary, not sufficient
  - stronger forms of invariance harder to motivate
- proof-theoretic approach:
  - rule format, harmony: unclear results, and unclear range of possible meanings
- Other criteria:
  - semantic completeness?
  - consistency with a consequence relation $\vdash$?

Keisler's $L(Q_1)$ shows that iso-invariance + completeness is not enough.

$\forall$ has non-standard interpretations consistent with $\models^{FO}$ (Church).

Suggests a mix: invariance + consistency (completeness doesn't help).

# Logicality and consequence

To what extent does a consequence relation in a language $L$ (a syntactic relation between sets of $L$-sentences and $L$-sentences) fix the meaning of certain symbols of $L$?

Should one even require logical symbols to have the property of being completely determined in this way?

Rudolf Carnap, in his 1943 book *The Formalization of Logic*, thought so.

The book tries to state and resolve the worry that this seems to fail even for classical propositional logic $CL$, i.e. that $\vdash^{CL}$ doesn't fix the meanings (truth tables) of the standard connectives.

# Carnap's Problem

The argument: A valuation is a bivalent assignment of T,F to all formulas.

Define the valuation $V^*$ by: $V^*(\varphi) = \mathrm{T}$ iff $\varphi$ is a tautology.

$V^*$ is consistent with classical propositional consequence $\vdash^{CL}$.

But $V^*(p) = V^*(\neg p) = \mathrm{F}$, and $V^*(p \vee \neg p) = \mathrm{T}$: not the standard truth table for disjunction!

(The table for $\wedge$, on the other hand, is fixed by $\vdash^{CL}$; a lack of symmetry that troubled Carnap.)

# Formal semantics to the rescue

If we apply the perspective of formal semantics (from Montague on), Carnap's Problem for $CL$ is solved.

In particular, valuations of proposition letters should extend compositionally to complex formulas.

The valuation $V^*$ is not compositional!

**Fact**

*If compositionality is required, $\vdash^{CL}$ determines the standard interpretations (truth tables) of all the connectives.*

(In fact, the proof shows that they are already determined by the intuitionistic part $\vdash^{IL}$ of propositional logic.)

# Carnap's question in general, and for *FO* in particular

We can ask 'Carnap's question' about any logical language. For example, about *FO*. Then some things must be in place:

- A recursive (generative) syntax: for *FO*, as usual.
- A semantic framework, in which interpretations assign suitable semantic values to all primitive expressions: as usual, but also connectives and quantifiers must be interpreted.
- A 'truth definition', that compositionally extends interpretations to complex expressions: as usual for *FO*; in particular, the semantic values of formulas are sets of assignments.
- A notion of what it means for such an interpretation to be consistent with a consequence relation $\vdash$: obvious for *FO*.
- For each primitive expression, a precise range of possible interpretations: in particular, for $\forall$, this is the class of type $\langle 1 \rangle$ generalized quantifiers, i.e. on each domain $M$, the set of subsets of $M$.
- For each putative logical expression, a notion of a standard interpretation (?): the standard interpretation of $\forall$ on $M$ is $\{M\}$.

# First-order logic

Propositional connectives must be standard; the question is about $\forall$.

**Theorem (Bonnay & W-hl 2016)**

*The interpretation $I$ over a domain $M$ is consistent with $\models^{FO}$ iff $I(\forall)$ is a principal filter on $\mathcal{P}(M)$. Moreover, if $I(\forall)$ is required to be permutation invariant ('topic-neutral'), then $I(\forall) = \{M\}$. Also, in this case, $I(=)$ is standard identity.*

The property of $FO$ that forces the filter $\mathcal{Q} = I(\forall)$ to be principal is:

$$\vdash^{FO} \forall x \forall y \varphi \leftrightarrow \forall y \forall x \varphi$$

which says that for all $R \subseteq M^2$, and with $R_a = \{b\colon aRb\}$,

$$\{a\colon R_a \in \mathcal{Q}\} \in \mathcal{Q} \;\Leftrightarrow\; \{a\colon (R^{-1})_a \in \mathcal{Q}\} \in \mathcal{Q} \quad \text{(commutativity)}$$

**Proof.**

Note that the Fréchet filter is non-commutative, and generalize to all non-principal filters. □

# Caveat

Assume that

- the language has predicate variables (or restrict the result to definable sets);
- the language has at least one binary predicate variable (results for the monadic case not known);
- the language does not have constant or function symbols (results extend by adding closure conditions).

# $L(Q_1)$

When $\mathcal{Q}$ is a principal filter generated by a set $A$,

$$\mathcal{M} \models \forall x \varphi [f] \Leftrightarrow A \subseteq [\![\varphi]\!]_{\mathcal{M},f}$$

which is 'outer domain' semantics in free logic.

So (consistency with) $\models^{FO}$ doesn't fix the meaning of $\forall$, but combined with iso-invariance it does.

Contrast with $L(Q_1)$ which is iso-invariant, axiomatizable, but the intended meaning of $Q_1$ cannot be recovered from $\models^{L(Q_1)}$:

Theorem (Bonnay & Speitel)

$(\mathcal{M}, Q_M)$ *is consistent with* $\models^{L(Q_1)}$ *iff* $Q_M = \{X \subseteq M : |X| \geq \aleph_\alpha\}$ *for some regular cardinal* $\aleph_\alpha$. *(NB Here* $\forall$ *and* $\exists$ *have their standard meaning.)*

**Conjecture:** If the interpretation of $Q$ is fixed (assuming iso-invariance) by the corresponding consequence relation $\models^{L(Q)}$, then $Q$ is $FO$-definable.

# Modal logic: semantic framework

Syntax:

$$p \mid \neg\varphi \mid (\varphi \wedge \psi) \mid \Box\varphi$$

Semantics: possible worlds semantics: given a set $W$ (worlds, states, points), the semantic values of sentences are subsets of $W$.

It follows from this, by compositionality, that the connectives must be interpreted (in $W$) as functions on sets of worlds (of appropriate arity).

So the range of possible interpretations is well defined.

Is there a standard interpretation for each connective?

For $\neg$ and $\wedge$ the answer is clear: complement and intersection.

What about $\Box$?

# Modal logic as applied

$\Box$ must be interpreted as a function from $\mathcal{P}(W)$ to $\mathcal{P}(W)$; which one?

Try: the function that expresses necessity as truth in all worlds. That is,

$$(F_{\text{uni}})^W(\Box)(X) = \begin{cases} W & \text{if } X = W \\ \emptyset & \text{otherwise} \end{cases}$$

This is an application of modal logic: $\Box$ as metaphysical necessity.

The machinery of modal logic is used today to model many other things: knowledge, belief, obligation, time, tense, provability, the dynamics of belief revision or information update, program execution, games, etc. etc.

These applications come with many different logics/consequence relations!

Very different from *FO*, and no obvious standard interpretation.

# $\neg$ and $\wedge$ are standard

In principle, an interpretation (over $W$) is a triple

$$F^W = \langle F^W(\neg), F^W(\wedge), F^W(\Box)\rangle$$

where $F^W(\neg), F^W(\Box)\colon \mathcal{P}(W) \to \mathcal{P}(W)$, and $F^W(\wedge)\colon \mathcal{P}(W)^2 \to \mathcal{P}(W)$.

NB The values of $p, q, \ldots \in \mathit{Prop}$ are given by valuations $V\colon \mathit{Prop} \to \mathcal{P}(W)$.

A modal logic is a set $L$ of sentences in the basic modal language closed under classical tautological consequence.

**Fact (Bonnay & W-hl (2016))**
*Any interpretation consistent with a modal logic must interpret $\neg$ and $\wedge$ (hence $\vee$ and $\to$) standardly: over a set $W$, as complement and intersection, respectively.*

We assume consistency with $CL$, which means that only $\Box$ remains to interpret. So in what follows, 'interpretation' means 'interpretation of $\Box$'.

And we write $F^W(X)$ instead of $F^W(\Box)(X)$.

# Local and global interpretations

A meaning of $\Box$ isn't tied to a universe $W$; it has $W$ as a parameter.

**Definition**

A simple local interpretation over $W$ is a function $F^W : \mathcal{P}(W) \to \mathcal{P}(W)$. A simple global interpretation $F$ associates with each $W$ a simple local interpretation $F^W$ over $W$.

$F_{\text{uni}}$ above is a simple global interpretation. So is $F_{\text{id}}$, defined by

$$(F_{\text{id}})^W(X) = X$$

# Semantic values and truth

A simple local interpretation $F^W$, together with a valuation $V$ on $W$, that is, a model $\mathcal{M} = (W, F^W, V)$, gives a semantic value $[\![\varphi]\!]_\mathcal{M} \subseteq W$ to each sentence $\varphi$:

- $[\![p]\!]_\mathcal{M} = V(p)$
- $[\![\neg\varphi]\!]_\mathcal{M} = W - [\![\varphi]\!]_\mathcal{M}$
- $[\![\varphi \wedge \psi]\!]_\mathcal{M} = [\![\varphi]\!]_\mathcal{M} \cap [\![\psi]\!]_\mathcal{M}$
- $[\![\Box\varphi]\!]_\mathcal{M} = F^W([\![\varphi]\!]_\mathcal{M})$

We can write

$$\mathcal{M}, w \models \varphi$$

instead of $w \in [\![\varphi]\!]_\mathcal{M}$.

It's an ordinary truth definition, where $\Box$ means whatever $F$ says it means.

# Local interpretations are neighborhood frames

A function $F^W : \mathcal{P}(W) \to \mathcal{P}(W)$ can also be presented as, for each $w \in W$, a family $F^{W,\Box,w}$ of subsets of $W$:

$$F^{W,\Box,w} = \{X \subseteq W : w \in F^W(X)\}$$

$F^{W,\Box,w}$ specifies for which $\varphi$ the formula $\Box\varphi$ is true at $w$.

We can recover $F^W$ from $\{F^{W,\Box,w}\}_{w \in W}$ and vice versa; these are just two perspectives on the same interpretation of $\Box$ over $W$.

Structures of the form $\mathcal{G} = (W, \{F^{W,\Box,w}\}_{w \in W})$, or equivalently, $(W, F^W)$, are known as neighborhood frames.

The notion of truth in a frame + valuation at a world in neighborhood semantics is the same as ours:

$$(\mathcal{G}, V), w \models \Box\varphi \quad \text{iff} \quad [\![\varphi]\!]_{(\mathcal{G},V)} \in F^{W,\Box,w} \quad \text{iff} \quad w \in F^W([\![\varphi]\!]_{(\mathcal{G},V)})$$

# Neighborhood semantics

In general, there are no restrictions on the families of subsets of $W$ in a neighborhood frame, just as any function $F^W\colon \mathcal{P}(W) \to \mathcal{P}(W)$ is allowed.

Validity in all neighborhood frames is axiomatized by the system $\mathbf{E} = CL +$ a single rule:

(E) If $\vdash \varphi \leftrightarrow \psi$, then $\vdash \Box\varphi \leftrightarrow \Box\psi$.

(E) just says that $F^W$ is extensional (a function on sets).

Neighborhood semantics generalizes Kripke semantics: Kripke frames are essentially neighborhood frames where each $F^{W,\Box,w}$ is a principal filter.

Validity in these frames is axiomatized by $\mathbf{K}$.

Neighborhood semantics is the most general possible worlds semantics.

# Consistency with a logic (consequence relation)

Back to Carnap's question: we are interested in how consequence relations constrain the interpretation of $\Box$, i.e. in results of the form:

$F$ is consistent with the logic $L$ iff $F$ has property P.

If $\mathcal{C} = \{F \colon F \text{ is P}\} = \{F_{\text{st}}\}$, then $L$ completely fixes the interpretation of $\Box$.

**Definition**

$F^W$ is consistent with a logic $L$ if $\vdash^L \varphi$ (i.e. $\varphi \in L$) implies that

for all valuations $V$, $[\![\varphi]\!]_{(W,F^W,V)} = W$

A simple global interpretation $F$ is consistent with $L$ if each $F^W$ is. (We consider validity rather than consequence — no harm ensues.)

# Consistency with **K**, S4, S5 translates to properties of $F$

**Fact**

(a) *$F^W$ is consistent with* **K** *iff* $F^W(W) = W$ *and* $F^W(X \cap Y) = F^W(X) \cap F^W(Y)$ *for all* $X, Y \subseteq W$.

(b) *$F^W$ is consistent with S4 iff, in addition, we have* $F^W(X) \subseteq X$ *and* $F^W(X) \subseteq F^W(F^W(X))$ *for all* $X \subseteq W$.

(c) *$F^W$ is consistent with S5 iff, in addition to the requirements for* **K** *we have* $F^W(X) \subseteq X$ *and* $W - F^W(X) \subseteq F^W(W - F^W(X))$ *for all* $X \subseteq W$.

Or, in terms of the families $F^{W,\Box,w}$:

**Fact**

(a) *$F^W$ is consistent with* **K** *iff each* $F^{W,\Box,w}$ *is a (not necessarily proper) filter.*

(b) *$F^W$ is consistent with S4 iff the following holds, for all* $w \in W$*:*
   1. $F^{W,\Box,w}$ *is a non-empty filter.*
   2. $w \in \bigcap F^{W,\Box,w}$
   3. *If* $X \in F^{W,\Box,w}$*, then* $F^W(X) \in F^{W,\Box,w}$.

(c) *$F^W$ is consistent with S5 iff, in addition to the requirements in (b), we have:*
   4. *If* $X \notin F^{W,\Box,w}$*, then* $W - F^W(X) \in F^{W,\Box,w}$.

# The local nature of truth

So far we have no standard interpretation of $\Box$.

And so far we have said nothing about what many think is the characteristic trait of modal logic: the local nature of truth:

> For the truth/falsity of $\varphi$ at $w$, only the worlds accessible from $w$ matter.

(Locality in this sense is not part of neighborhood semantics.)

Ways to remedy this:

(i) Let an accessibility relation $R$ be a parameter of interpretations.

(ii) Consider simple local interpretations/neighborhood frames that do respect the local nature of truth.

Start with (ii).

# Kripkean interpretations

**Definition**

The associated accessibility relation to $F^W$, $Acc_{F^W}$, is defined by:

$$w \, Acc_{F^W} \, w' \quad \text{iff} \quad w' \in \bigcap F^{W,\Box,w}$$

I.e. $w'$ is accessible from $w$ precisely when it belongs to all (semantic values of) $\varphi$ such that $\Box\varphi$ is true at $w$ according to $F^W$.

**Definition**

$F^W$ is Kripkean if

$$F^W(X) = \{w \in W : (Acc_{F^W})_w \subseteq X\}$$

**Fact**

*Tfae:*

(a) *$F^W$ is Kripkean / generated by $Acc_{F^W}$.*

(b) *Each $F^{W,\Box,w}$ is a principal filter.*

# Standard interpretations again

It makes sense to call Kripkean interpretations standard (if you think the local nature of truth is essential).

Then the ‘Carnapian’ task would be to see if some modal logic determines precisely these.

To answer this, it’s useful to step back and adopt a slightly more abstract perspective.

# The standard Galois connection

Abstractly: $L$ is a logic, $\mathcal{C}$ a class of frames (e.g. Kripke or neighborhood). Validity on a frame, $F \models \varphi$, is as usual.

(Leave out $^W$ for simplicity.)

**Definition**

(a) $Val(L) = \{F \colon \forall \varphi \in L \;\; F \models \varphi\}$

(b) $Log(\mathcal{C}) = \{\varphi \colon \forall F \in \mathcal{C} \;\; F \models \varphi\}$

Then we have the usual (antitone) Galois connection:

(GC) $\mathcal{C} \subseteq Val(L) \;\Leftrightarrow\; L \subseteq Log(\mathcal{C})$

# Completeness and correspondence

$$\mathcal{C} \subseteq Val(L) \Leftrightarrow L \subseteq Log(\mathcal{C})$$

**Definition**

(a) $L$ corresponds to $\mathcal{C}$ if $\mathcal{C} = Val(L)$.

(b) $L$ is (sound and) complete wrt $\mathcal{C}$ if $L = Log(\mathcal{C})$.

(c) $L$ is complete if $L$ is complete wrt some $\mathcal{C}$.

We can reconstruct what Carnap was looking for in *The Formalization of Logic* to be correspondence, in addition to completeness.

> *PC does not completely fulfill its purpose; it is not a full formalization of propositional logic.*
>
> *Thus the rules of PC, both in proofs and derivations, yield all those and only results for which they were made. What else could we require of them?*
>
> *The conclusion . . . that PC is a complete formal representation of . . . the logical properties of the propositional connectives . . . is wrong. This is shown by the possibility of non-normal interpretations.* (pp. 96–97)

# Examples

The situation is very familiar in modal logic. Let $\mathcal{EQ}$ be the class of equivalence relations:

(1) $\mathcal{EQ} = \mathit{Val}(S5)$ (correspondence)

(2) $S5 = \mathit{Log}(\mathcal{EQ})$ (completeness)

Also, if $\mathcal{U}$ is the class of universal Kripke frames, $(W, W^2)$, we have

(3) $S5 = \mathit{Log}(\mathcal{U})$

which, since $\mathcal{U} \subsetneq \mathcal{EQ}$, shows that the property of being universal is not modally definable.

# Partition interpretations and S5

For neighborhood frames, let $\mathcal{KR}$ be the class of Kripkean interpretations, and $\mathcal{PI}$ the subclass of partition interpretations: $F$ is Kripkean and $Acc_F$ is an equivalence relation.

**Theorem**

(a) $\mathcal{PI} = Val(S5)$ *(correspondence)*

(b) $S5 = Log(\mathcal{PI})$ *(completeness)*

**Proof.**

(b) follows easily from the completeness wrt Kripke frames. (a) requires a new proof. □

**Fact**

*The class $\mathcal{KR}$ of Kripkean interpretations is not modally definable.*

**Proof.**

It is known that $\mathbf{K} = Log(\mathcal{KR}) = Log(\mathcal{FI})$, where $\mathcal{FI}$ is the class of frames $(W, F^W)$ s.t. each $F^{W,\Box,w}$ is a filter. But $\mathcal{KR} \subsetneq \mathcal{FI}$. □

# Consistency and correspondence

Recall that we wanted results of the form:

$F$ is consistent with $L$ iff $F \in \mathcal{C}$

This is equivalent to

$\mathcal{C} = Val(L)$

So it is really correspondence results we are after, but

- we look at $F$ as an interpretation of $\Box$;
- we'd like $\mathcal{C}$ to be as small as possible (ideally a unit set).

# Topo-interpretations

A topology on $W$ is a set $\tau$ of subsets of $W$ (called the open sets) containing $\emptyset, W$ and closed under finite intersections and arbitrary unions. $A \subseteq W$ is closed if $W - A$ is open. The interior of $A$ is the largest open subset of $A$:

$$int_\tau(A) = \bigcup\{X \in \tau : X \subseteq A\}$$

Definition

$F^W$ is a topo-interpretation if $F^W = int_\tau$ for some topology $\tau$ on $W$.

Let $\mathcal{TP}$ be the class of topo-interpretations.

Theorem

(a) $\mathcal{TP} = Val(S4)$ *(correspondence)*

(b) $S4 = Log(\mathcal{TP})$ *(completeness, McKinsey and Tarski, 1944)*

Proof.

(a) is also known; in fact, if $F^W$ is consistent with S4, then

$$\tau = \{F^W(X) : X \subseteq W\}$$

is a topology such that $F^W = int_\tau$. □

# A topological proof of the correspondence result for S5

We must show that if $F^W$ is consistent with S5, $F^W$ is a partition interpretation.

The 'hard' part is to show that $F^W$ is Kripkean.

Let $\tau = \{F(X) \colon X \subseteq W\}$, so that $F^W = int_\tau$ (consistency with S4).

It is easy to see that consistency with S5 (the B axiom) implies that $\tau$ is locally indiscrete: every open set is closed. That is,

(1) $X \in \tau \;\Leftrightarrow\; W - X \in \tau$

Let $R = Acc_{F^W}$; we must show that $w \in F(X) \Leftrightarrow R_w \subseteq X$.

If $w \in F(X)$ it follows by the definition of $Acc_{F^W}$ that $R_w \subseteq X$.

If $w \notin F(X)$, there is, by (1), $Z$ such that $F(Z) = W - F(X)$. Using this, and the fact that $Acc_{F^W}$ can be shown to be the so-called specialization pre-order of $\tau$,

$$wRw' \;\Leftrightarrow\; \forall Y \in \tau (w \in Y \Rightarrow w' \in Y)$$

it is not hard to verify that $R_w \not\subseteq X$.

# Approximating standard interpretations

No logic corresponds to the class of standard/Kripkean interpretations.

But only standard/Kripkean interpretations are consistent with S5.

This is far from the case for S4. Define the simple interpretation $F_{\mathrm{Fr}}$ (where 'Fr' is for 'Fréchet') by

$$(F_{\mathrm{Fr}})^W(X) = \begin{cases} X & \text{if } X \text{ is co-finite} \\ \emptyset & \text{otherwise} \end{cases}$$

$F_{\mathrm{Fr}}$ is consistent with S4 (in fact, with S4.3), but not with S5.

For infinite $W$, each filter $F_{\mathrm{Fr}}^{W,\Box,w}$ is non-principal, and $\bigcap F_{\mathrm{Fr}}^{W,\Box,w} = \{w\} \notin F_{\mathrm{Fr}}^{W,\Box,w}$.

$Acc_{(F_{\mathrm{Fr}})^W}$ is the identity relation, same as $Acc_{(F_{\mathrm{id}})^W}$, but $(F_{\mathrm{Fr}})^W$ is not standard with respect to $Acc_{(F_{\mathrm{Fr}})^W}$, so $F_{\mathrm{Fr}}$ is not Kripkean.

**Conjecture:** S5 is minimal (at least among normal extensions of S4.3) with this property.

# $\Box$ versus $\forall$

Compare the results about these two:

- In both cases, the 'hard' part was to show that the interpretation must be a principal filter.

But the methods differ: for $\forall$ the crucial fact was

(1) $\vdash^{FO} \forall x \forall y \varphi \leftrightarrow \forall y \forall x \varphi$

This essentially uses the 2-variable fragment of *FO*, which is not within the bounded fragment, corresponding to modal logic.

**K** allows many non-standard interpretations. We can add axioms to avoid some of them, but S4 is not enough.

Why is S5, that is,

(2) $\vdash^{S5} \varphi \rightarrow \Box\Diamond\varphi$

enough?

# Permutation invariance for $\forall$ and $\Box$

Permutation invariance is reasonable for $\forall$, but too strong for $\Box$ in general — cf. the existence of standard interpretations.

But we can get a Carnap style result for iso-invariant interpretations of $\Box$.

Call a simple global interpretation $F$ trivial if for some $W$ with $|W| \geq 2$, $F^W = (F_{\text{id}})^W$ or $F^W = (F_{\text{emp}})^W$.

**Theorem**

*The only non-trivial permutation invariant simple global interpretation consistent with S5 is $F_{uni}$.*

That is, under permutation invariance, metaphysical necessity is pinned down by S5.

But in general we should ask for invariance under modal transformations: iso-invariance wrt $Acc_{F^W}$, or weaker, such as p-morphisms.

# Unique standard interpretations?

Meanings are global: compare (generalized) quantifiers: The meaning of, say, $\exists$ is not tied to a particular domain or universe; it is not, for example, the set of non-empty subsets of $\mathbb{N} = \{0, 1, 2, \ldots\}$.

Rather, it is the function mapping any universe $M$ to the set of non-empty subsets of $M$.

The global $F_{\text{uni}}$ captures (one version of) $\Box$ as metaphysical necessity.

No logic determines exactly the standard/Kripkean interpretations.

If we want one standard interpretation, we might try making the accessibility relation a parameter.

# Parametric interpretations

**Definition**

A parametric interpretation $I$ maps each Kripke frame $(W, R)$ to a simple local interpretation, written $I^{(W,R)}$, of $\Box$.

There is now a unique standard parametric interpretation of $\Box$:

$$I_{\text{st}}^{(W,R)}(X) = \{w \in W \colon R_w \subseteq X\}$$

This gives the usual clause for $\Box\varphi$:

$$(W, R, V), w \models \Box\varphi \text{ iff for all } w' \text{ s.t. } wRw', (W, R, V), w' \models \varphi$$

One can consider other forms of dependence on $R$:

$$I_{\text{tr}}^{(W,R)}(X) = \{w \in W \colon (R^*)_w \subseteq X\} \qquad (^* \text{ is transitive closure})$$

Or no dependence at all:

$$I_{\text{uni}}^{(W,R)} = (F_{\text{uni}})^W$$

# Familiar invariances

We can now consider non-local properties of interpretations, i.e. properties that relate the behavior of $I$ on different frames:

- invariance under generated subframes
- iso-invariance
- p-morphism invariance

**Fact**

*P-morphism invariance entails iso-invariance and invariance under generated subframes.*

Which logic characterizes $I_{\mathrm{st}}$?

# A Carnap style characterization of $I_{\text{st}}$

We cannot simply say that a parametric interpretation $I$ is consistent with a logic $L$ if each $I^{(W,R)}$ is consistent with $L$.

Then $I_{\text{st}}$ would not be consistent with S5, unless $R$ is an equivalence relation.

Instead, we relativize to a class $\mathcal{C}$ of frames:

**Definition**

$I$ is $\mathcal{C}$-consistent with $L$ if whenever $\vdash^L \varphi$, we have, for all $(W, R) \in \mathcal{C}$ and valuations $V$ over $W$, $[\![\varphi]\!]_{(W, I^{(W,R)}, V)} = W$.

Again, say that $I$ is trivial if for some $W$ with $|W| \geq 2$, $I^{(W,R)} = I_{\text{id}}^{(W,R)}$.

**Theorem**

*Suppose $I$ is a non-trivial parametric interpretation which is $\mathcal{EQ}$-consistent with S5 and invariant under automorphisms and generated subframes. Then for each $(W, R) \in \mathcal{EQ}$, $I^{(W,R)} = I_{st}^{(W,R)}$ (and hence $R = Acc_{I^{(W,R)}}$).*

# Global properties wrt $Acc_{F^W}$

But recall that each interpretation $F^W$ has an associated accessibility relation $Acc_{F^W}$ (even if it is not always a 'real' accessibility relation).

So, forgetting about unique standard interpretations, we can consider global properties, like iso- and generated subframe and p-morphism invariance, for simple (not parametric) global interpretations, wrt their associated accessibility relations.

For example, $F_{\text{uni}}$ is then (trivially) invariant under generated subframes, in contrast with the parametric $I_{\text{uni}}$ (defined by $I_{\text{uni}}^{(W,R)}(X) = F_{\text{uni}}(X)$).

But $F_{\text{Fr}}$ is not.

**Fact**

*Kripkean simple global interpretations are p-morphism invariant (hence invariant under isomorphisms and generated subframes) in this sense.*

# Characterizing Kripkean interpretations, again

In view of the preceding Fact, wouldn't it be nice if Kripkean/standard interpretations are precisely the ones invariant under p-morphisms?

Unfortunately, that's not the case ...

It's easy to find p-morphism invariant but non-Kripkean interpretations not consistent with **K**; for example, $F$ defined by $F^W(X) = \emptyset$ for all $X \subseteq W$.

[Since $F^W(W) \neq W$, Necessitation fails, so $F$ is inconsistent with **K**, and (hence) not Kripkean, but trivially p-morphism invariant.]

More seriously, there are also p-morphism invariant but non-Kripkean interpretations consistent with **K**.

# A counter-example

Take a Kripke frame $(\mathbb{N}, R)$ such that the only p-morphism from $(\mathbb{N}, R)$ to $(\mathbb{N}, R)$ is the identity (such frames exist).

Define $F^{\mathbb{N}}$ by 'Fréchetization':

$$F^{\mathbb{N}}(X) = \begin{cases} \{w \in \mathbb{N} \colon R_w \subseteq X\} & \text{if } X \text{ is co-finite} \\ \emptyset & \text{otherwise} \end{cases}$$

Since $\bigcap\{X \subseteq \mathbb{N} \colon R_w \subseteq X \;\&\; X \text{ is co-finite}\} = R_w$, we have $Acc_{F^{\mathbb{N}}} = R$.

So the only p-morphism from $(\mathbb{N}, Acc_{F^{\mathbb{N}}})$ to $(\mathbb{N}, Acc_{F^{\mathbb{N}}})$ is the identity.

$F^{\mathbb{N}}$ is not standard wrt $R$, hence not Kripkean, but consistent with **K**.

For $W \neq \mathbb{N}$, let $F^W(X) = W$ for $X \subseteq W$, so $Acc_{F^W} = \emptyset$.

We obtain a simple global interpretation $F$ which is non-Kripkean, p-morphism invariant, and consistent with **K**.

It would be nice to know more about p-morphism invariant interpretations consistent with **K**.

# Conclusion

The mix: Consistency with a logic/consequence relation + Invariance:

- a neat alternative to strengthening invariance;
- new model-theoretic questions;
- nice model-theoretic results for *FO* and extensions;
- a wilder landscape for ML.