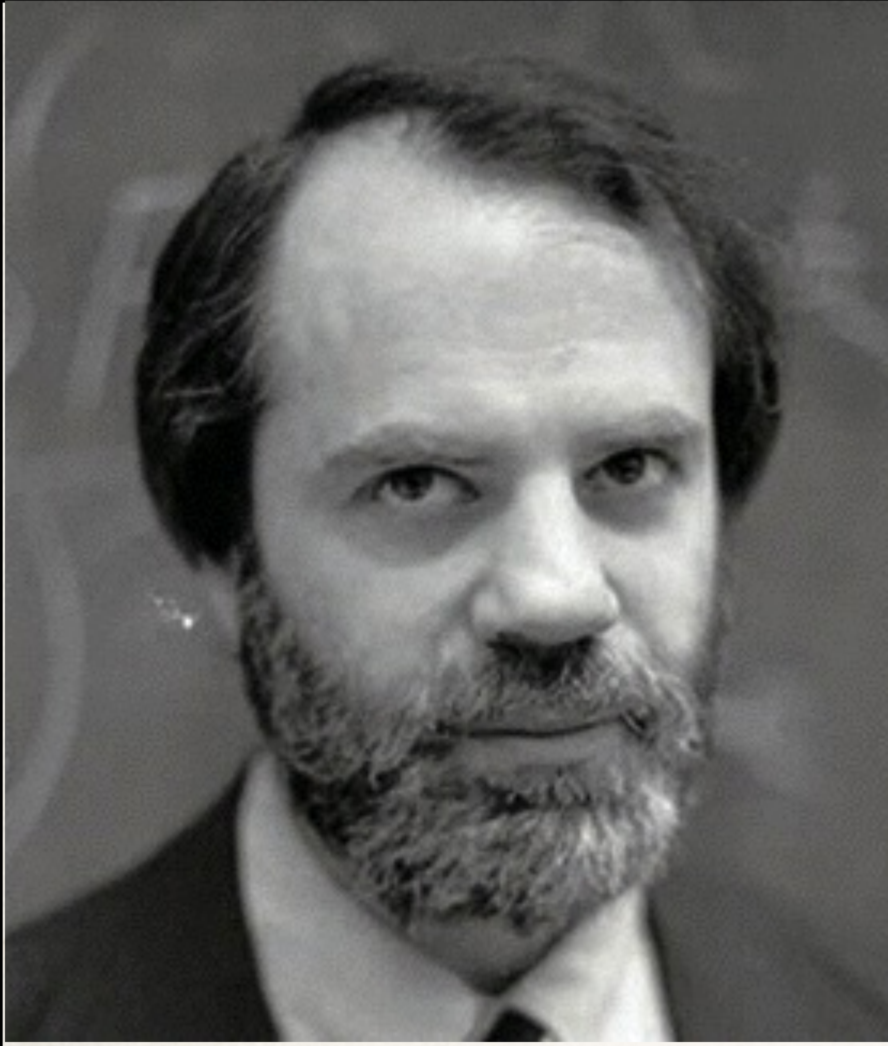
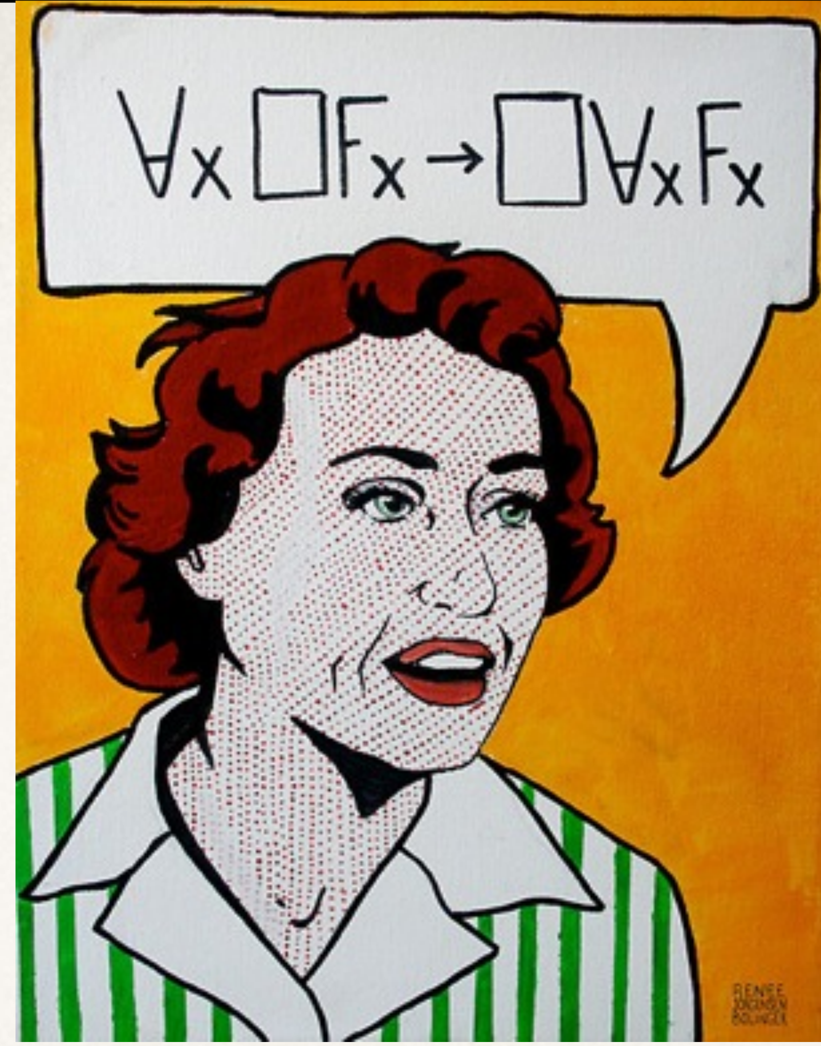
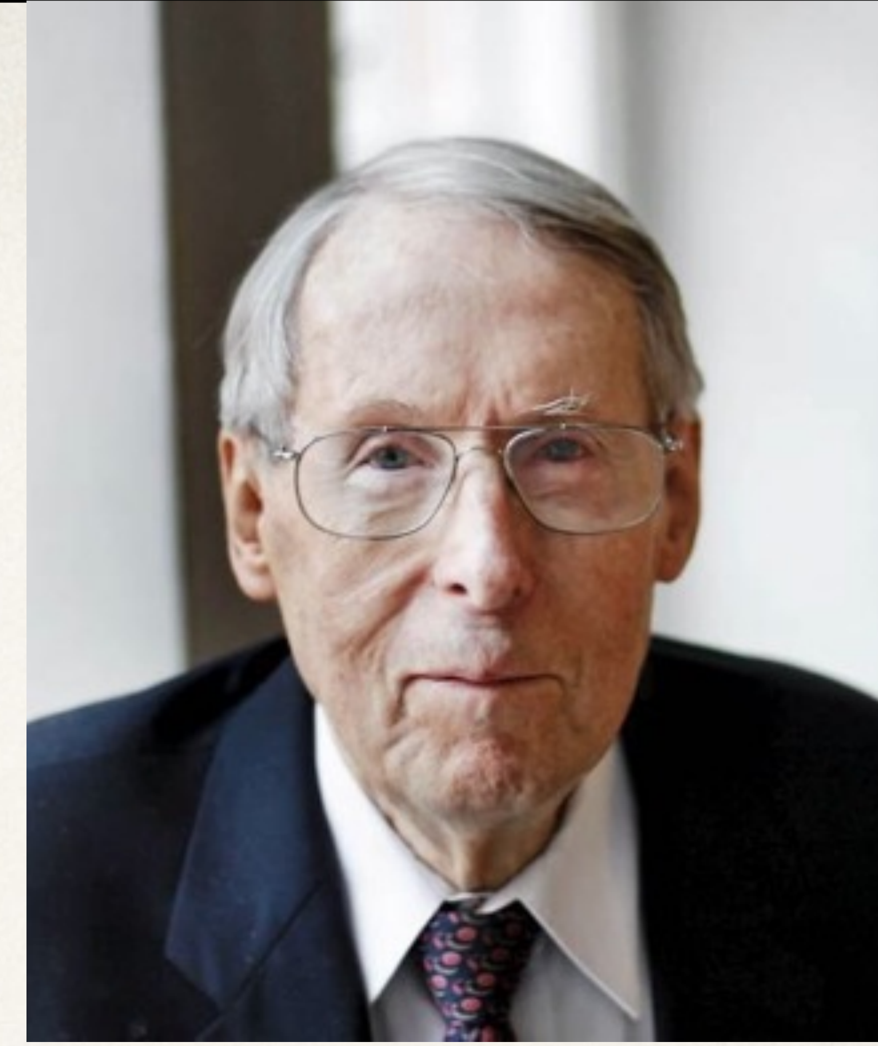


Saul Kripke



Ruth Barcan Marcus



Jaakko Hintikka

PHIL 50 - Introduction to Logic

Marcello Di Bello, Stanford University, Spring 2014

Week 9 — Wednesday Class - Modal Logic

Statements of Modal Logic

Statements in Propositional and Predicate Logic are typically **about the world as it is**

Statements in Modal Logic are **about the world as it could be** (e.g. the world in our imagination, the world in our beliefs, the world in the future or in the past)

Statements in Modal Logic:

"Possibly (A) "
"Necessarily (A)"

A Key Notion in Modal Logic is **POSSIBLE WORLD**

What is a possible world?

Two Views About Possible Worlds

Robust view of possible worlds

Possible worlds are alternative worlds made up of things, objects, people, etc. which **exist somewhere far away** from the actual (or current) world.

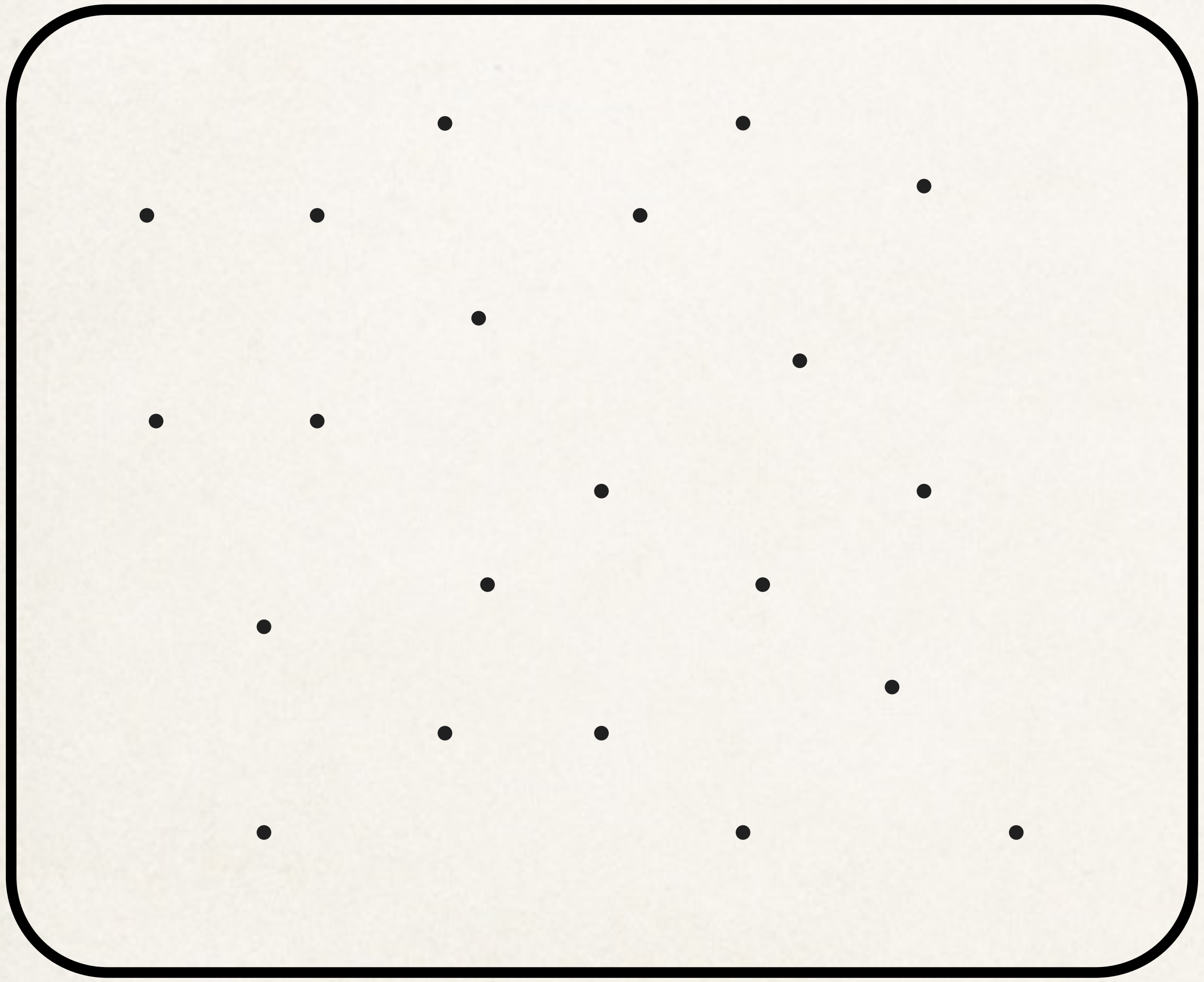
This position resembles the doctrine of "parallel universes".

Thin view of possible worlds

Possible worlds are alternatives to the actual world but they **need not exist somewhere far away** from the actual world.

We Shall Not Take Side in the Dispute Between the Robust and the Thin View of Possible Worlds

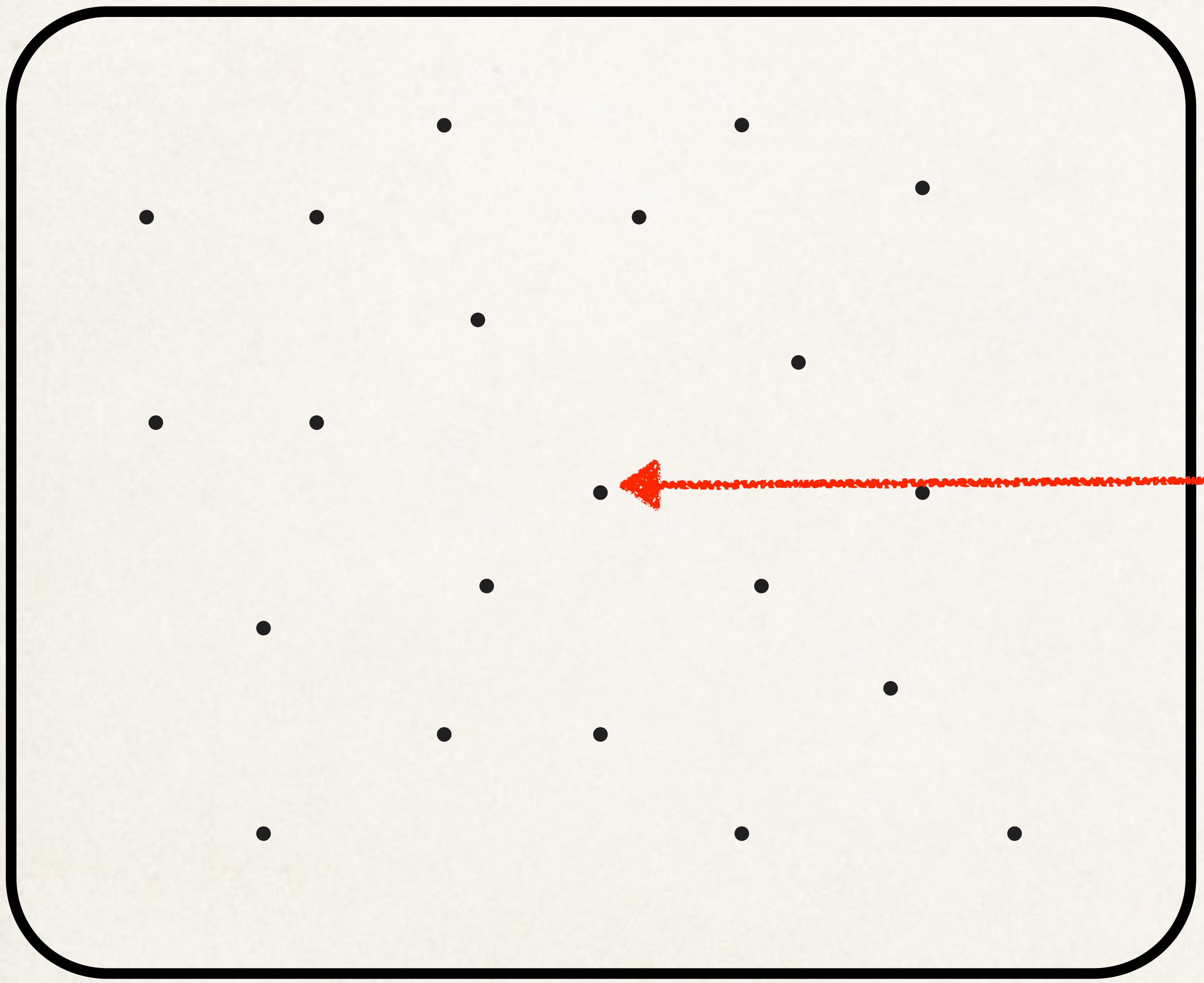
W — The Set of all Possible Worlds



We can represent a possible world simply as a dot.

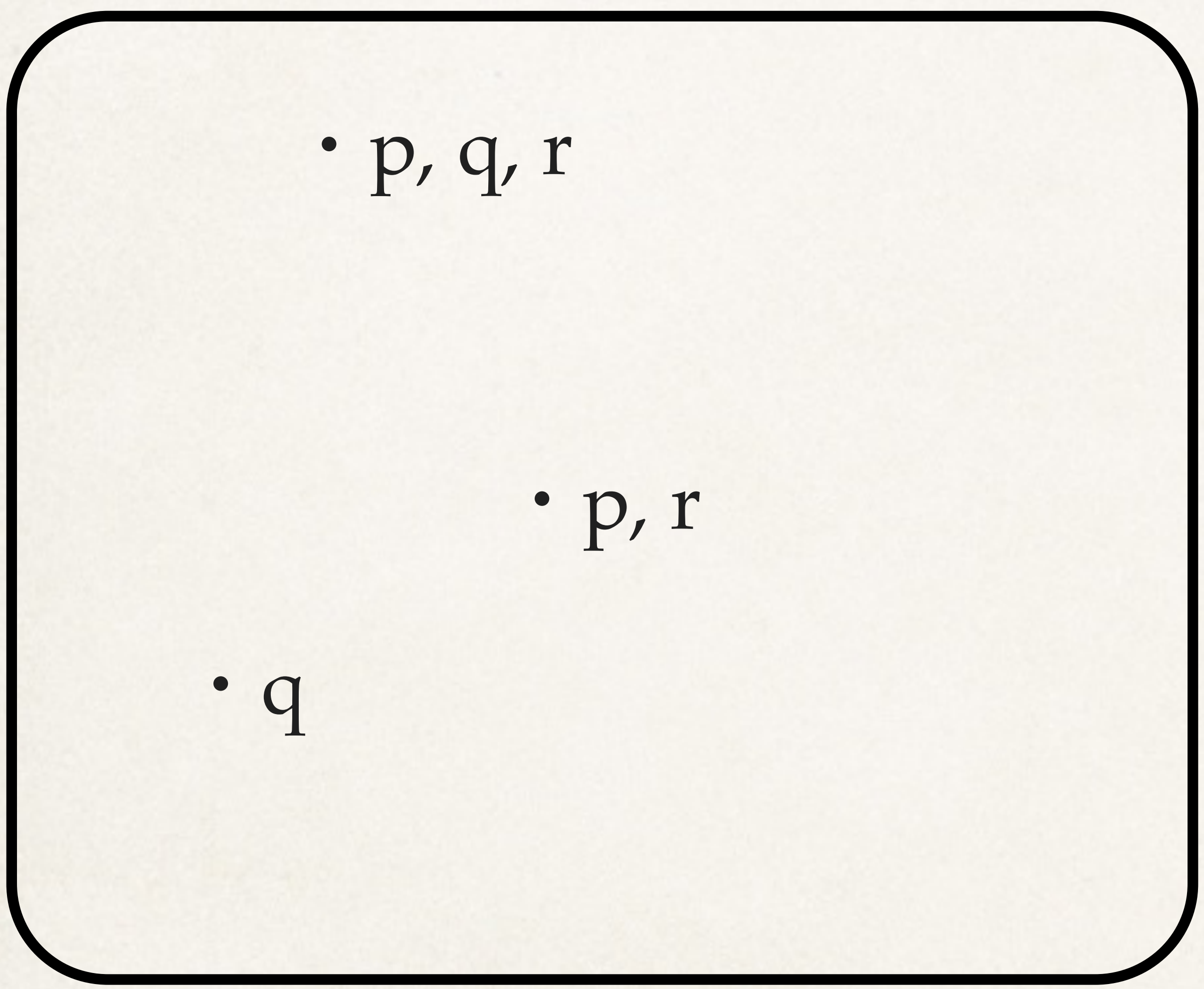
We can represent the set ***W*** of all possible worlds as a diagram containing many (possibly infinite) dots.

The Actual World — the World We Live in



We can single out the actual world from the other possible worlds.

Different Formulas are True at Different Worlds

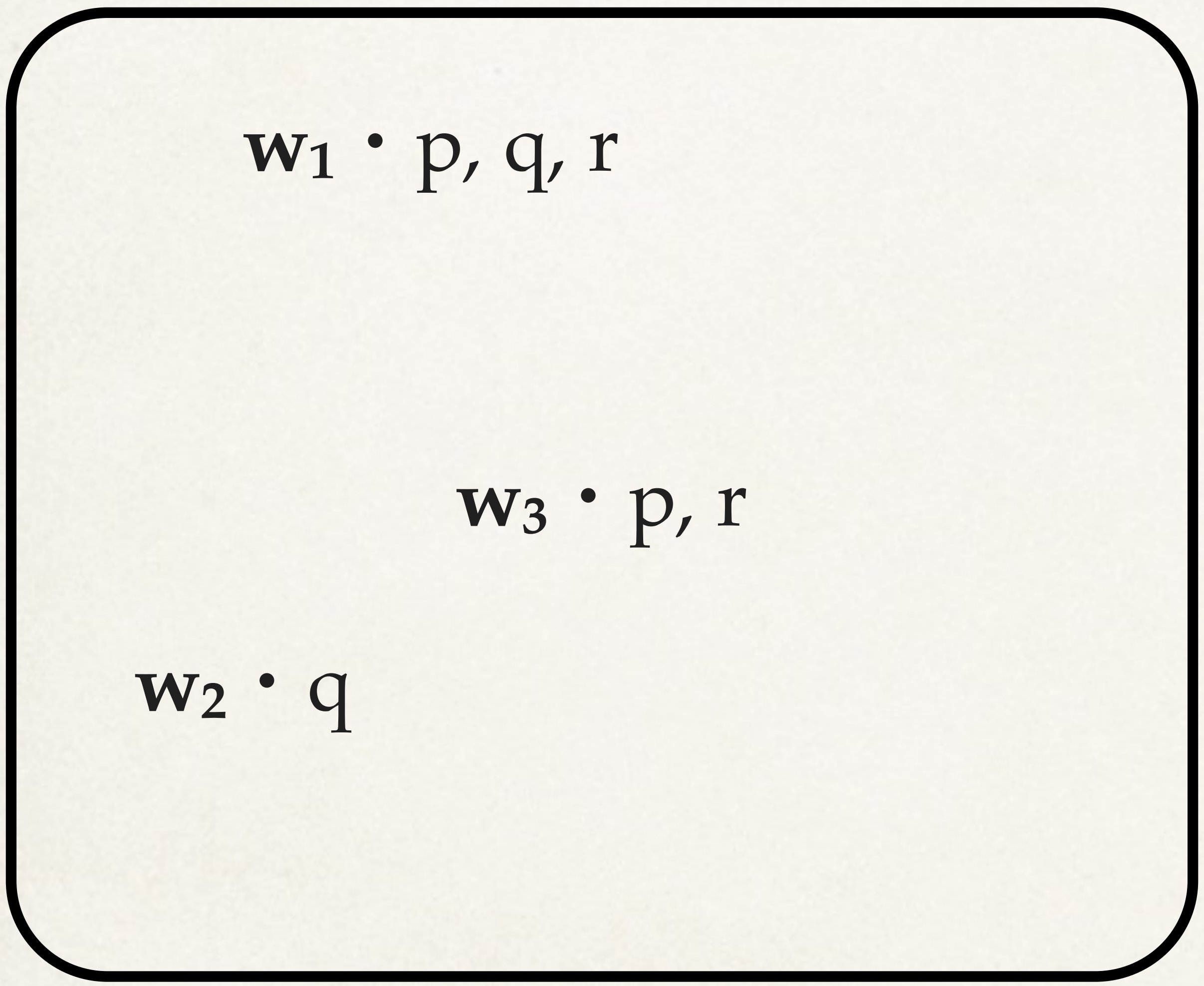


In a world, formulas **p**, **q**, and **r** are all true.

In another world, formula **q** is true while formulas **p** and **r** are false.

In still another world, formulas **r** and **p** are true and **q** is false.

We Can Label Possible Worlds



Formulas **p**, **q**, and **r** are all true in possible world $\mathbf{w_1}$

Formula **q** is true and formulas **p** and **r** are false in possible world $\mathbf{w_2}$

Formulas **r** and **p** are true and formula **q** is false in possible world $\mathbf{w_3}$

Notation:

$\mathbf{w} \vDash \phi$
iff
ϕ is true at world $\mathbf{w}$

$\mathbf{w} \nvDash \phi$
iff
ϕ is false at world $\mathbf{w}$

$\mathbf{w_1}$ • p, q, r

$\mathbf{w_3}$ • p, r

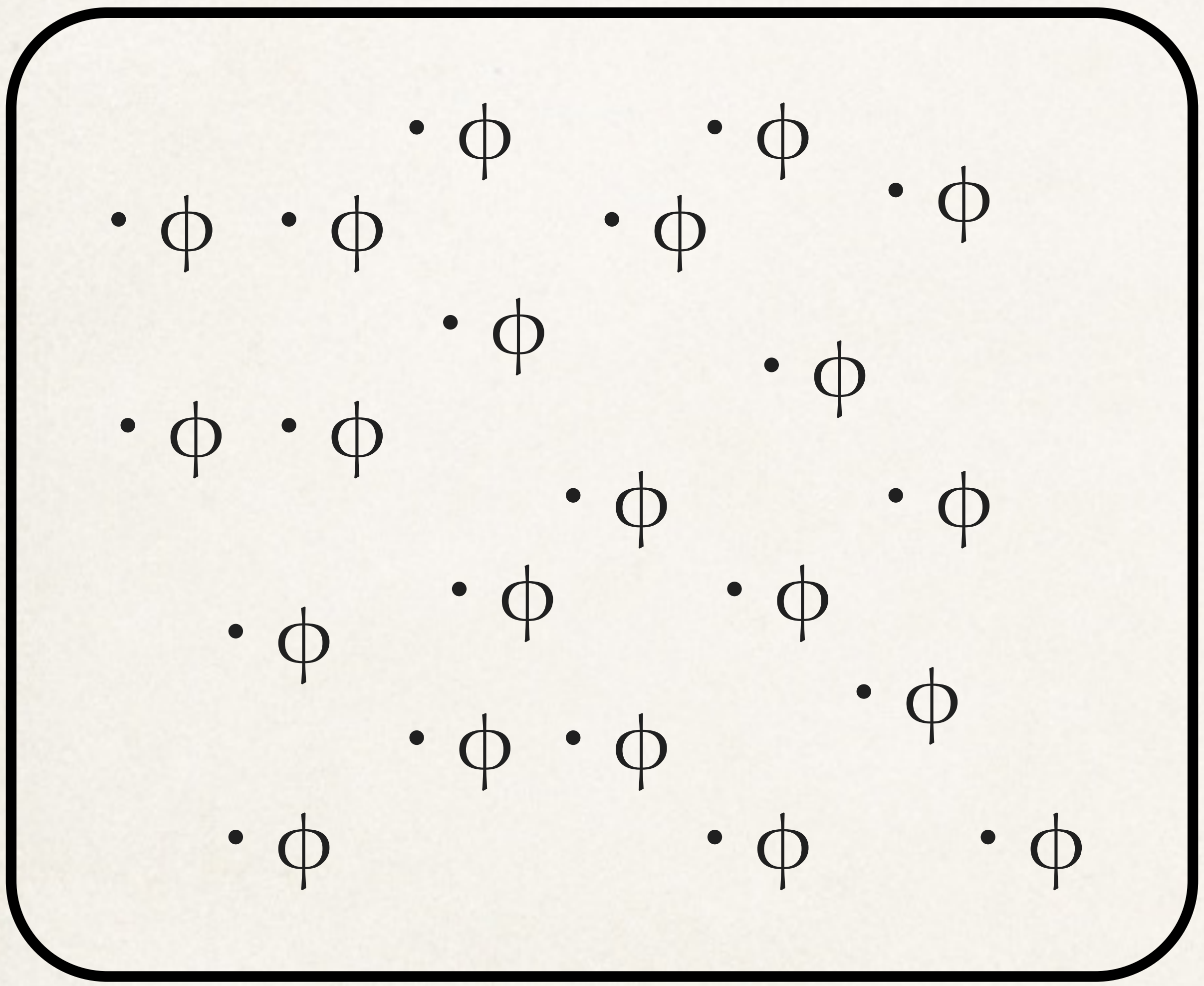
$\mathbf{w_2}$ • q

$\mathbf{w_1} \vDash p$
$\mathbf{w_1} \vDash q$
$\mathbf{w_1} \vDash r$

$\mathbf{w_2} \nvDash p$
$\mathbf{w_2} \vDash q$
$\mathbf{w_2} \nvDash r$

$\mathbf{w_3} \vDash p$
$\mathbf{w_3} \nvDash q$
$\mathbf{w_3} \vDash r$

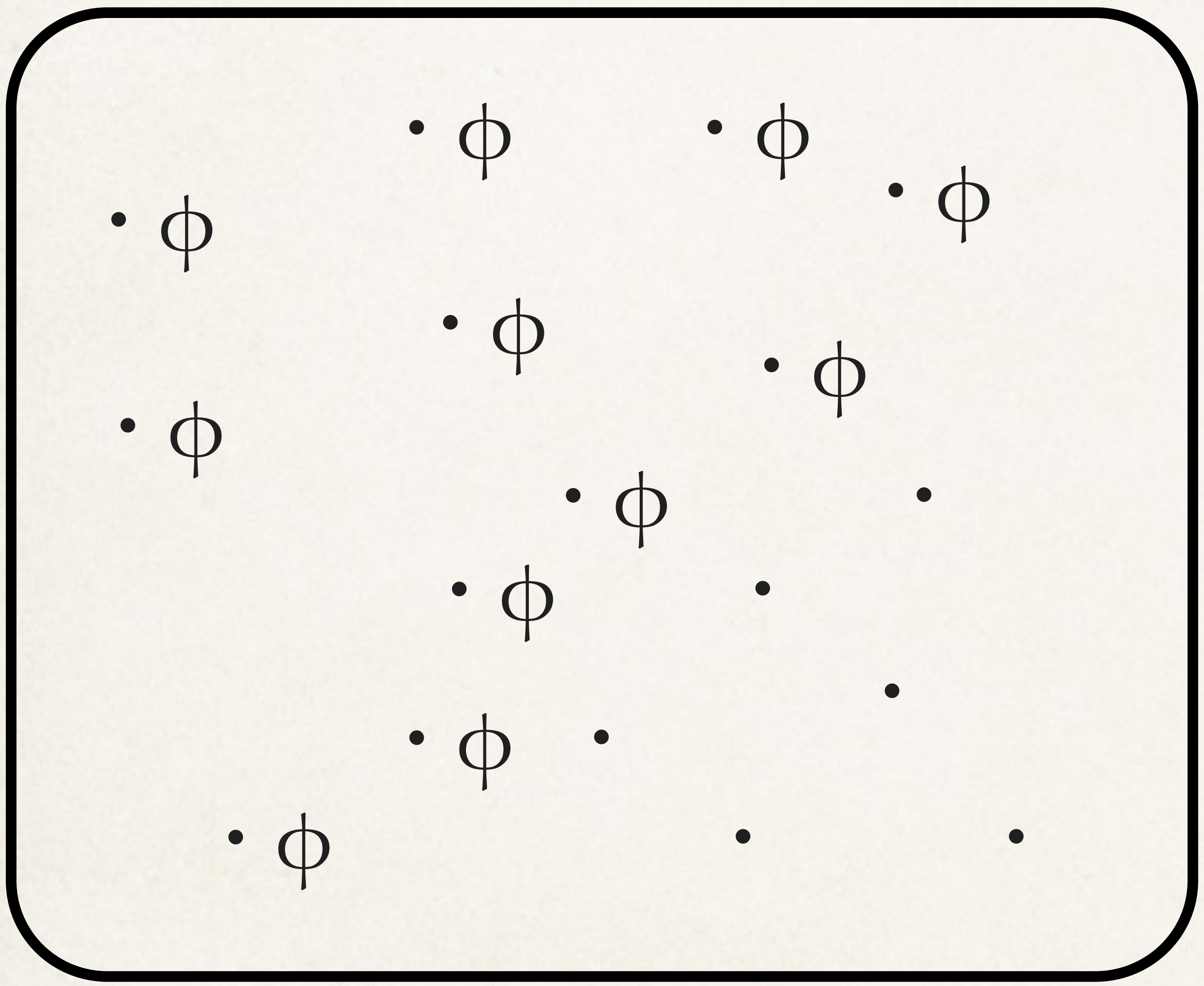
What if a Formula is True in All Possible Worlds?



Formula **ϕ** is true in all possible worlds, so it is ***necessarily true***

To express that a formula is ***necessarily true,*** we shall write □**ϕ** (read "**box ϕ**")

What if a Formula is True in Some (Or At Least One) Possible World(s)?



Formula **ϕ** is true in some possible worlds, so it is ***possibly true***

To express that a formula is ***possibly true***, we shall write ◇**ϕ** (read "**diamond ϕ**")

NB: *A formula can be both necessarily true and possibly true.*

We Now Enrich Our Modal Logic with Predicates, Constants, and Quantifiers

Inside a Possible World

For each world **w**, we have:

- domain $\mathbf{D_w}$ of objects
- interpretation function $\mathbf{I_w}$ for constant symbols and predicate symbols
- a variable assignment $\mathbf{g_w}$ for variable symbols

NB: $\mathbf{D_w}$, $\mathbf{I_w}$, $\mathbf{g_w}$ can be different from one world to another

*In **predicate logic**, the truth of a formula is relative to a fixed model **M**.*

*In **modal (predicate) logic** the truth of a formula is relative to each possible world and each possible world is similar to a model of predicate logic.*

A Model **M** in Modal Logic

*In **predicate logic**, a model consists of*

*(1) a domain **D**,*
*(2) an interpretation **I**, and*
*(3) an assignment function **g**.*

*There are many models, depending on how **D**, **I**, and **g** are defined.*

*In **predicate modal logic**, a model is a set **W** of possible worlds **w**,* where *each world **w** consists of*

(1) a domain $\mathbf{D}_w$,
(2) an interpretation $\mathbf{I}_w$, and
(3) an assignment function $\mathbf{g}_w$.

*There are many models, depending on how $\mathbf{D}_w$, $\mathbf{I}_w$, and $\mathbf{g}_w$ are defined for each possible world **w**.*

Truth in a Possible World

$\mathbf{M}, w \vDash \mathbf{P}(c)$ *iff* $\mathbf{I_w(c)} \in \mathbf{I_w(P)}$

$\mathbf{M}, w \vDash a=b$ *iff* $\langle \mathbf{I_w}(a), \mathbf{I_w}(b)\rangle\ \mathbf{I_w(=)}$

$\mathbf{M}, w \vDash \neg\, \phi$ *iff* $\mathbf{M}, w \nvDash \phi$

$\mathbf{M}, w \vDash \phi \wedge \psi$ *iff* $\mathbf{M}, w \vDash \phi$ ***and*** $\mathbf{M}, w \vDash \psi$

$\mathbf{M}, w \vDash \phi \vee \psi$ *iff* $M, w \vDash \phi$ ***or*** $\mathbf{M}, w \vDash \psi$

$\mathbf{M}, w \vDash \phi \rightarrow \psi$ *iff* $M, w \vDash \phi$ ***implies*** $\mathbf{M}, w \vDash \psi$

$\mathbf{M}, w \vDash \exists x\phi$ *iff* there is a d [$d \in D_w$ and $\langle D_w, I_w, g_{w[x:=d]}\rangle \vDash \phi$]

$\mathbf{M}, w \vDash \forall x\phi$ *iff* for all d [*if* $d \in D_w$, *then* $\langle D_w, I_w, g_{w[x:=d]}\rangle \vDash \phi$]

Note how truth conditions are indexed to the possible world w

…and the Truth Conditions for the Modal Formulas…

$\mathbf{M}, w \vDash \Box\, \phi$ *iff* for all v, it holds that $\mathbf{M}, v \vDash \phi$

$\mathbf{M}, w \vDash \Diamond\, \phi$ *iff* for some v, it holds that $\mathbf{M}, v \vDash \phi$

Notation for Truth and Validity

TRUTH in a model:

Instead of writing

$$\mathbf{M}, w \vDash \phi$$

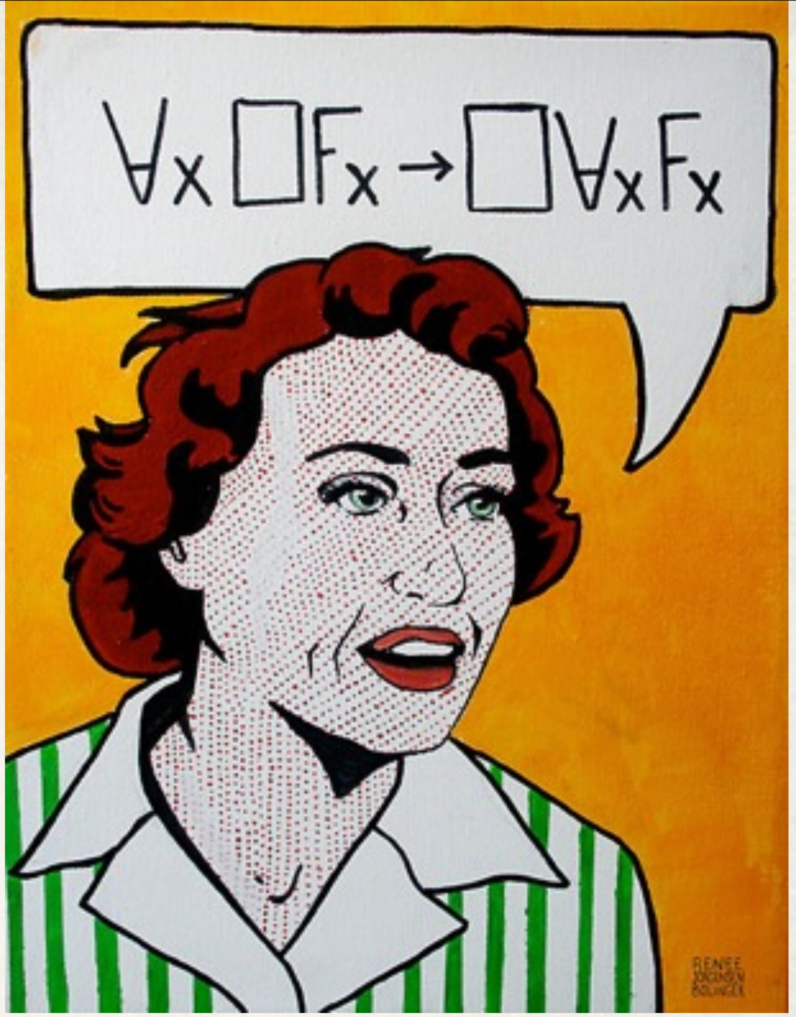
we can also simply writing

$$w \vDash \phi$$

VALIDITY:

$\vDash \phi$ *iff* $M, w \vDash \phi$ for all models M and all possible worlds w.

Let’s Discuss a Controversial Formula by Ruth Barcan Marcus



The Barcan Formula: $\forall x \Box Px \rightarrow \Box \forall x Px$

If everything is necessarily **P**, then necessarily everything is **P**.

Barcan Formula: ∀x□Px→□∀xPx

It is Valid? *Not Necessarily*

Consider a model **M** with two worlds *w* and *v* where:

D_w={❀, ☞} and I_w(P)={❀, ☞}

D_u={❦, ☞, ❀} and I_u(P)={❀, ☞}

g_w=g_u

Is the Barcan Formula always true?

No, because the antecedent is true in *w*, but the consequent is false in *w*.

$\forall x \Box Px \to \Box \forall x Px$ It is Valid? (1)

$D_w = \{$❀, ☞$\}$ and $I_w(P) = \{$❀, ☞$\}$

$D_u = \{$❦, ☞, ❀$\}$ and $I_u(P) = \{$❀, ☞$\}$

$\mathbf{M}, w \vDash \forall x \Box Px$ *iff*

for all **d** [*if* $d \in D_w$, *then* $\langle D_w, I_w, g_{w[x:=d]} \rangle \vDash \Box Px$]

for all **d** [*if* $d \in D_w$, *then for all worlds* v, $\langle D_v, I_v, g_{v[x:=d]} \rangle \vDash Px$]

for all **d** [*if* $d \in D_w$, *then for all worlds* v, $g_{v[x:=d]}(x) \in I_v(P)$]

for all **d** [*if* $d \in D_w$, *then for all worlds* v, $d \in I_v(P)$] (*)

Note that (*) holds because

(1) we should only consider the objects in D_w namely ❀ and ☞, and

(2) we should check that both objects ❀ and ☞, in all possible worlds, namely w and u, are in the sets $I_w(P)$ and $I_u(P)$.

Now, both ❀ and ☞ are in $I_w(P)$ and $I_u(P)$, so (*) holds.

So, $\mathbf{M}, w \vDash \forall x \Box Px$. ***<u>The antecedent is true in w</u>***.

$\forall x \Box Px \rightarrow \Box \forall x Px$ It is Valid? (2)

$D_w = \{$❀, ☞$\}$ and $I_w(P) = \{$❀, ☞$\}$

$D_u = \{$❦, ☞, ❀$\}$ and $I_u(P) = \{$❀, ☞$\}$

M, $w \vDash \Box \forall x Px$ *iff*

for all worlds v, it holds that $M, v \vDash \forall x Px$

for all worlds v, for all d [*if* $d \in D_v$, *then* $\langle D_v, I_v, g_{v[x:=d]} \rangle \vDash Px$]

for all worlds v, for all d [*if* $d \in D_v$, *then* $g_{v[x:=d]}\rangle(x) \in I_v(P)$]

for all worlds v, for all d [*if* $d \in D_v$, $d \in I_v(P)$] (**)

Note that (**) does hold because

(1) we should consider the objects in D_w and in D_u namely ❀, ☞, and ❦

(2) we should make sure that for each possible world all the objects in it belong to the set corresponding to the interpretation of *P*.

Now, object ❦ is in D_u, but ❦ is not in $I_u(P)$, so (**) does not hold.

So, **M**, $w \nvDash \Box \forall x Px$. ***<u>The consequent is false in w</u>*.**

We have just shown that the Barcan Formula $\forall x \Box Px \rightarrow \Box \forall x Px$ is invalid, but…

What If Domains Of Objects Do Not Differ From One Possible World to Another?

The Barcan formula was to shown to be invalid because we imagined possible worlds with different domains of objects.

If possible worlds are assumed to have the same domain of objects, the Barcan formula is valid. ***This is left for you as an exercise.***

Here's the model we used:

$\mathbf{D_w}$={❀, ☞} and **$\mathbf{I_w(P)}$={❀, ☞}**

$\mathbf{D_v}$={❦, ☞, ❀} and **$\mathbf{I_v(P)}$={❀, ☞}**