CS 295A/395D: Artificial Intelligence

KT45

Prof. Emma Tosch

25 March 2022

The University of Vermont

Logistics

Reminder: **this is neither an online nor a hybrid course.**

- I've built leniency into the schedule and grading scheme.
- I will support remote attendance insofar as it is convenient for me to do so. ***Do not rely on this being a remote course.***
- If you cannot attend in person, be in touch with me or Michael and come to student hours.

Recall: Basic modal logic syntax

$$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \Box\varphi \mid \Diamond\varphi$$

p	$\neg p$	$q \wedge p$	$p \rightarrow q$	$p \vee q$
$\Box p$	$\Diamond\neg p$	$\Box(q \wedge p)$	$\Diamond(p \rightarrow q)$	$q \vee \Box p$
$\Box\Box p$	$\neg\Diamond p$	$\Box q \wedge \Diamond p$	$\Diamond(p \rightarrow \Box q)$	$\Box\Diamond(q \vee \Box p)$

Recall: Basic modal logic syntax

$$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \Box\varphi \mid \Diamond\varphi$$

p	$\neg p$	$q \wedge p$	$p \rightarrow q$	$p \vee q$
$\Box p$	$\Diamond\neg p$	$\Box(q \wedge p)$	$\Diamond(p \rightarrow q)$	$q \vee \Box p$
$\Box\Box p$	$\neg\Diamond p$	$\Box q \wedge \Diamond p$	$\Diamond(p \rightarrow \Box q)$	$\Box\Diamond(q \vee \Box p)$

Recall: Basic modal logic syntax

$$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \Box\varphi \mid \Diamond\varphi$$

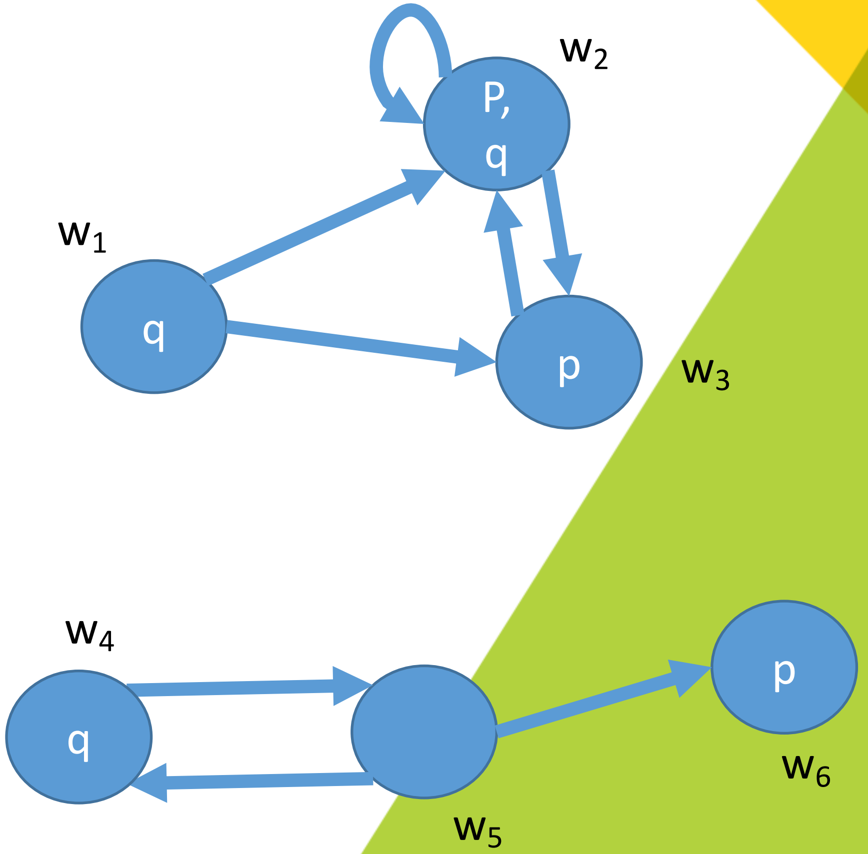
p	$\neg p$	$q \wedge p$	$p \rightarrow q$	$p \vee q$
$\Box p$	$\Diamond\neg p$	$\Box(q \wedge p)$	$\Diamond(p \rightarrow q)$	$q \vee \Box p$
$\Box\Box p$	$\neg\Diamond p$	$\Box q \wedge \Diamond p$	$\Diamond(p \rightarrow \Box q)$	$\Box\Diamond(q \vee \Box p)$

Recall: Basic modal logic semantics

Model structure: $\mathcal{M} = (W, R, L)$

- W: set of "worlds" (nodes in a graph),
- R: binary "accessibility relation" on W (edges in a graph),
- L: "labeling function" from each world to a subset of atoms,

Where the set of atoms is the set of propositions.



Basic modal logic: Semantics

$w \Vdash \top$

$w \nVdash \bot$

$w \Vdash$ a iff a ∈ L(w)

$w \Vdash \neg\varphi$ iff w $\nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff w $\Vdash \varphi$ and w $\Vdash \psi$

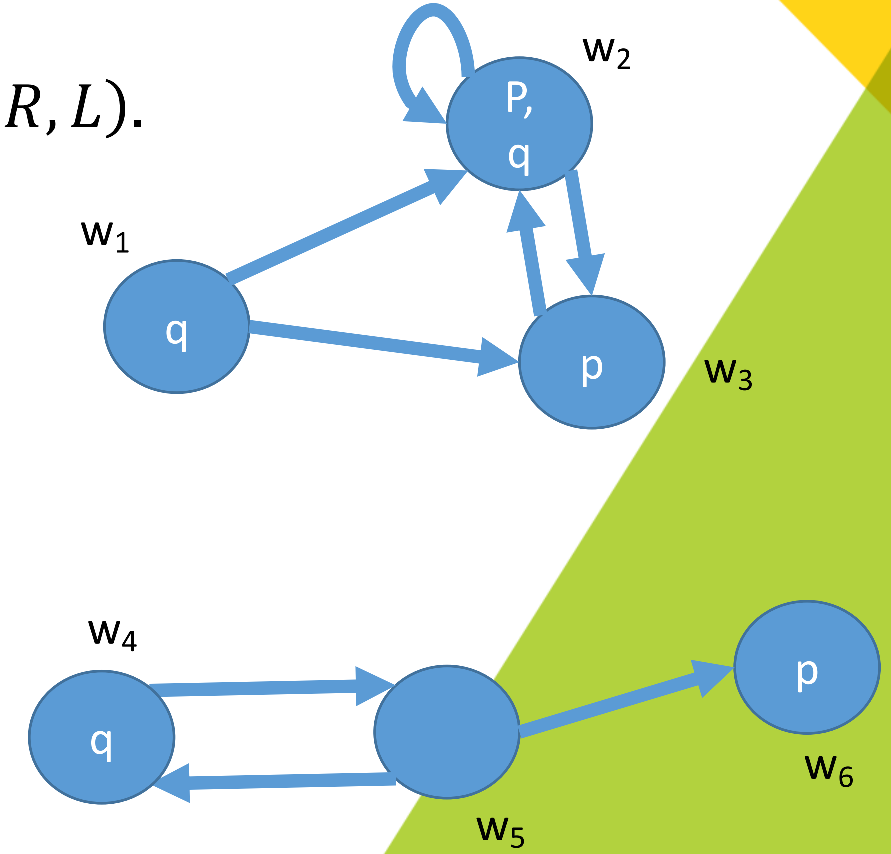
$w \Vdash \varphi \vee \psi$ iff at least one of w $\Vdash \varphi$ or w $\Vdash \psi$

$w \Vdash \varphi \rightarrow \psi$ iff w $\Vdash \varphi$ whenever w $\Vdash \psi$

$w \Vdash \Box\varphi$ iff ∀ w' ∈ W (R(w, w') → w' $\Vdash \psi$)

$w \Vdash \Diamond\varphi$ iff ∃ w' ∈ W (R(w, w') ∧ w' $\Vdash \psi$)

Given Model structure: $\mathcal{M} = (W, R, L)$.
Let $w \in W$.



Recap: New rules/equivalences

DeMorgan's:

$$\neg\Box\varphi = \Diamond\neg\varphi$$

$$\neg\Diamond\varphi = \Box\neg\varphi$$

Distributive:

$$\Box(\varphi \wedge \psi) = \Box\varphi \vee \Box\psi$$

$$\Diamond(\varphi \vee \psi) = \Diamond\varphi \wedge \Diamond\psi$$

Connective equivalence:

$$\neg\Box\neg\varphi = \Diamond\varphi$$

Tautology, contradiction:

$$\Box\top = \top$$

$$\Box\top \neq \Diamond\top$$

$$\Diamond\bot = \bot$$

$$\Diamond\bot \neq \Box\bot$$

Recap: Basic valid formula of modal logic

All the same valid formulas, plus "K":

$$(\Box(\varphi \rightarrow \psi) \wedge \Box\varphi) \rightarrow \Box\psi$$

Also written:

$$\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$$

Axiom	Knowledge	Belief	Property (axiom name)	R(w, w’)
$\Box p \rightarrow p$	Agent Q only knows true things	NOT A SUITABLE AXIOM	Reflexive (T)	$\forall\, w \in W\, (R(w, w))$ *Not true for belief!!!*
$\Box\varphi \rightarrow \Box\Box\varphi$	Agent Q knows what it knows (introspection)	Agent Q believes what it believes	Transitive (4)	$\forall (w, w', w'') \in (W \times W \times W), (R(w, w') \land R(w', w'') \rightarrow R(w, w''))$
$\diamond\, \varphi \rightarrow \Box \diamond \varphi$	If agent Q doesn’t know something, it doesn’t know what it doesn’t know	If agent Q doesn’t believe something, it doesn’t believe what it doesn’t believe	Euclidean (5)	$\forall (w, w', w'') \in (W \times W \times W), (R(w, w') \land R(w, w'') \rightarrow R(w', w''))$
$\diamond\, \top$	Agent Q doesn’t know contradictions	Agent Q doesn’t believe contradictions	Serial (D)	$\forall w \in W (\exists w' \in W (R(w, w'))$
$\Box\varphi \rightarrow \diamond\, \varphi$	Agent Q can chain knowledge (true things, doesn’t get stuck)	Agent Q can chain belief (true things, doesn’t get stuck)	Serial (D)	$\forall w \in W (\exists w' \in W (R(w, w'))$

Graph comparison on board

Natural deduction in modal logic

Recall: natural deduction is useful for deriving true things through syntactic manipulation alone.

All the same rules as propositional logic, plus:

1. Convert all $\Diamond\varphi$ to $\neg\Box\neg\varphi$
2. Open a new type of scope (dashed lines) for arbitrary possible world
3. Add relevant axioms (always K, sometimes others)

Epistemic logic engineering: S5/KT45^n

Give □ a new semantics: "Specialize" □ to mean knowledge *of a specific agent*

- New name: K (for knowledge, not for the formula schema K, named after Saul Kripke)
- "possible worlds" become "other agents'" knowledge:
  - $\Box p \equiv K_i\, p$: "Agent *i* knows *p*"
  - $\Box\Box p \equiv K_i K_j p$: "Agent *i* knows that agent *j* knows *p*"
  - $K_1 p \wedge K_2 p \wedge \cdots K_n p \equiv E_G p, G = \{1, \ldots, n\}$: "Every agent index by G knows *p*"
  - $E_G p \wedge E_G E_G p \wedge E_G E_G E_G p \;\wedge \cdots \equiv C_G$: "Everyone knows that everyone knows….. *p* (common knowledge)"
  - $D_G p$: "Knowledge of *p* is distributed among G" → *p* can be inferred from what G knows (**reachability**))

Basic modal logic: Semantics

$w \Vdash \top$

$w \nVdash \bot$

$w \Vdash$ a iff a ∈ L(w)

$w \Vdash \neg\varphi$ iff w $\nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff w $\Vdash \varphi$ and w $\Vdash \psi$

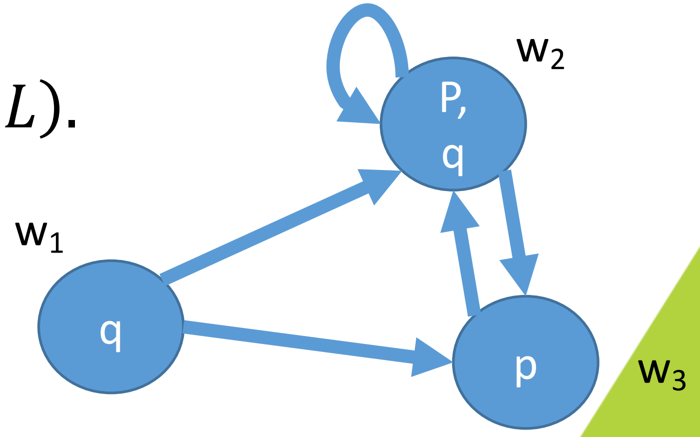
$w \Vdash \varphi \vee \psi$ iff at least one of w $\Vdash \varphi$ or w $\Vdash \psi$

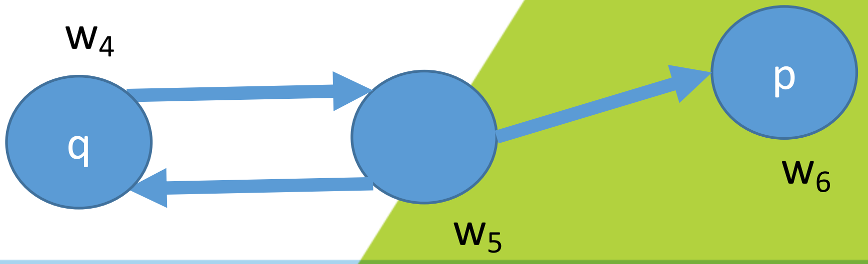
$w \Vdash \varphi \rightarrow \psi$ iff w $\Vdash \varphi$ whenever w $\Vdash \psi$

$w \Vdash \Box\varphi$ iff ∀ w' ∈ W (R(w, w') → w' $\Vdash \psi$)

$w \Vdash \Diamond\varphi$ iff ∃ w' ∈ W (R(w, w') ∧ w' $\Vdash \psi$)

Given Model structure: $\mathcal{M} = (W, R, L)$.
Let $w \in W$.





KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

$w \Vdash \top$

$w \nVdash \bot$

$w \Vdash$ a iff a ∈ L(w)

$w \Vdash \neg\varphi$ iff w $\nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff w $\Vdash \varphi$ and w $\Vdash \psi$

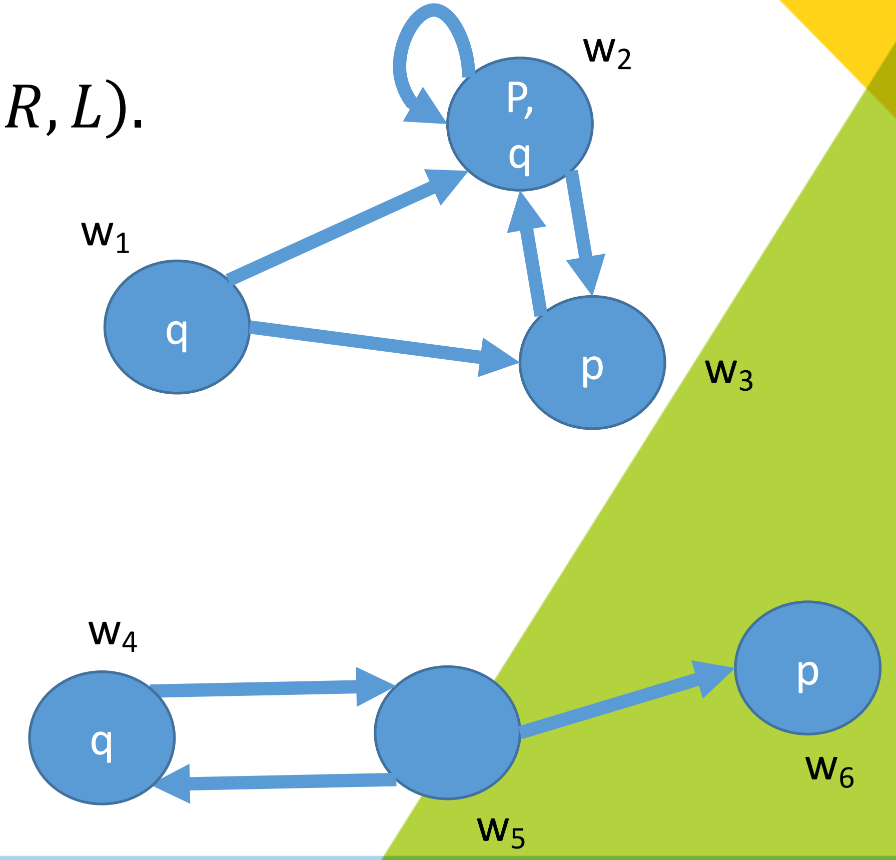
$w \Vdash \varphi \vee \psi$ iff at least one of w $\Vdash \varphi$ or w $\Vdash \psi$

$w \Vdash \varphi \rightarrow \psi$ iff w $\Vdash \varphi$ whenever w $\Vdash \psi$

$w \Vdash \Box\varphi$ iff ∀ w' ∈ W (R(w, w') → w' $\Vdash \psi$)

$w \Vdash \Diamond\varphi$ iff ∃ w' ∈ W (R(w, w') ∧ w' $\Vdash \psi$)

Given Model structure: $\mathcal{M} = (W, R, L)$.
Let $w \in W$.



KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i\in A\}}, L)$.

$w \Vdash \top$

$w \nVdash \bot$

$w \Vdash a$ iff $a \in L(w)$

$w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

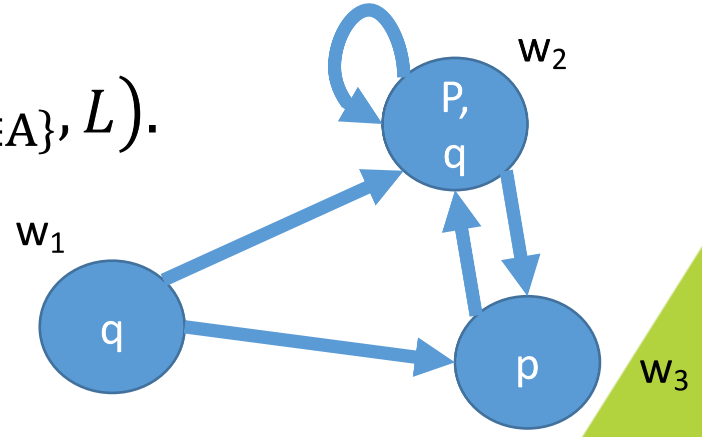
$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

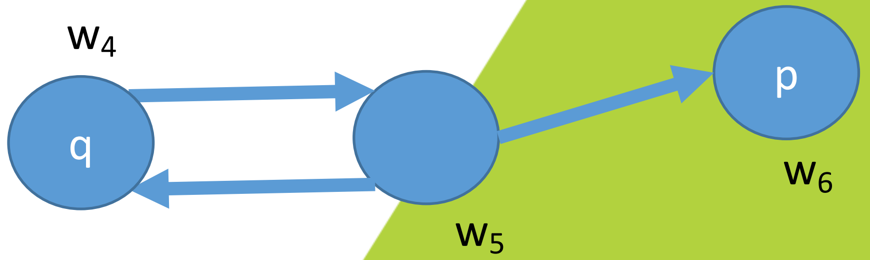
$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash \Box\varphi$ iff $\forall w' \in W\ (R(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash \Diamond\varphi$ iff $\exists w' \in W\ (R(w, w') \wedge w' \Vdash \psi)$





KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = \left(W, \{R\}_{\{i \in \mathrm{A}\}}, L\right)$.

$w \Vdash \top$

$w \nVdash \bot$

$w \Vdash$ a iff a $\in$ L(w)

$w \Vdash \neg\varphi$ iff w $\nVdash \varphi$

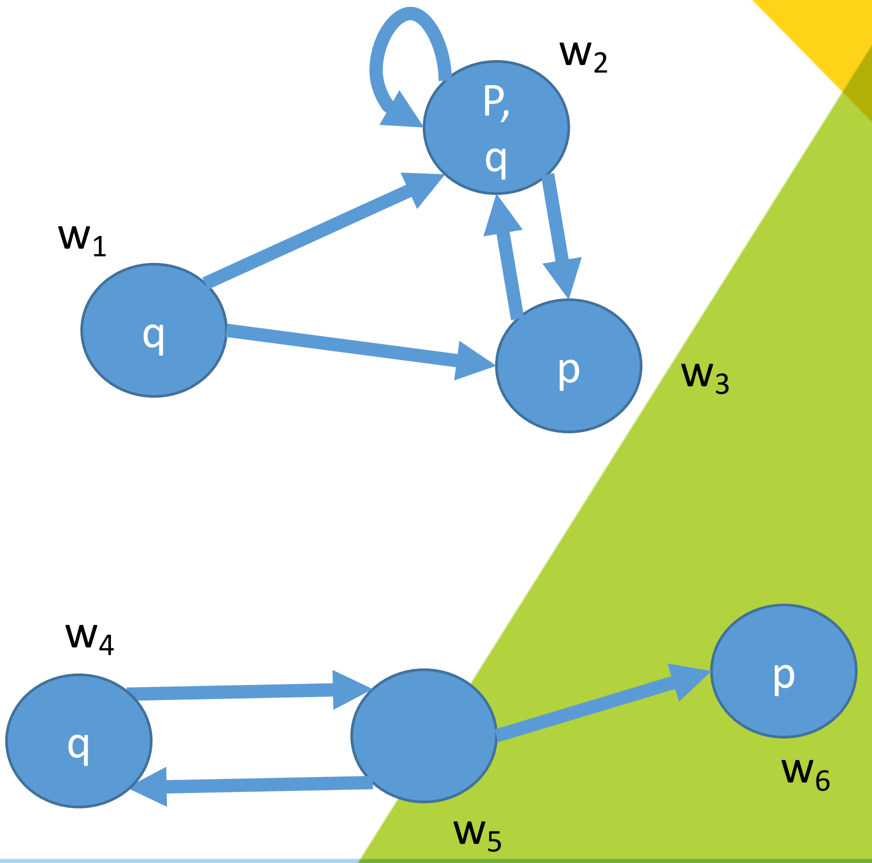
$w \Vdash \varphi \wedge \psi$ iff w $\Vdash \varphi$ and w $\Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of w $\Vdash \varphi$ or w $\Vdash \psi$

$w \Vdash \varphi \rightarrow \psi$ iff w $\Vdash \varphi$ whenever w $\Vdash \psi$

$w \Vdash \Box\varphi$ iff $\forall$ w' $\in$ W (R(w, w') $\rightarrow$ w' $\Vdash \psi$)

$w \Vdash \Diamond\varphi$ iff $\exists$ w' $\in$ W (R(w, w') $\wedge$ w' $\Vdash \psi$)



$\mathbf{KT45^n}$ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i\in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

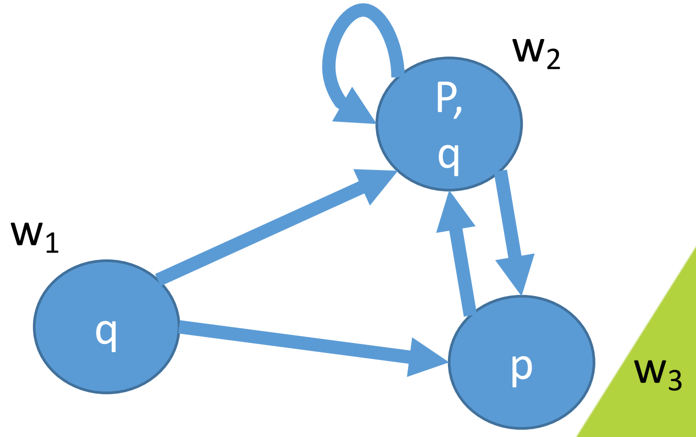
$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

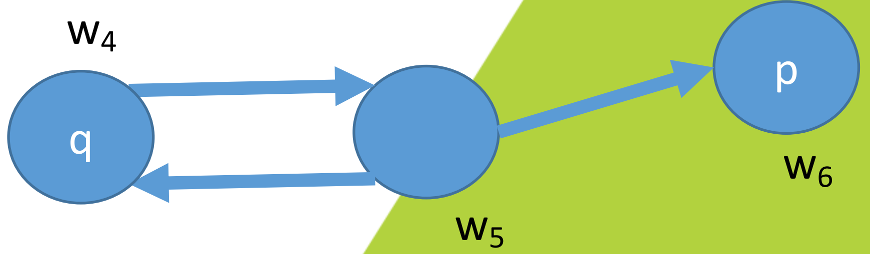
$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

$w \Vdash \varphi \to \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash \Box\varphi$ iff $\forall$ w' $\in$ W (R(w, w') $\to$ w' $\Vdash \psi$)

$w \Vdash \Diamond\varphi$ iff $\exists$ w' $\in$ W (R(w, w') $\wedge$ w' $\Vdash \psi$)





$\mathbf{KT45^n}$ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = \big(W, \{R\}_{\{i \in \mathrm{A}\}}, L\big)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

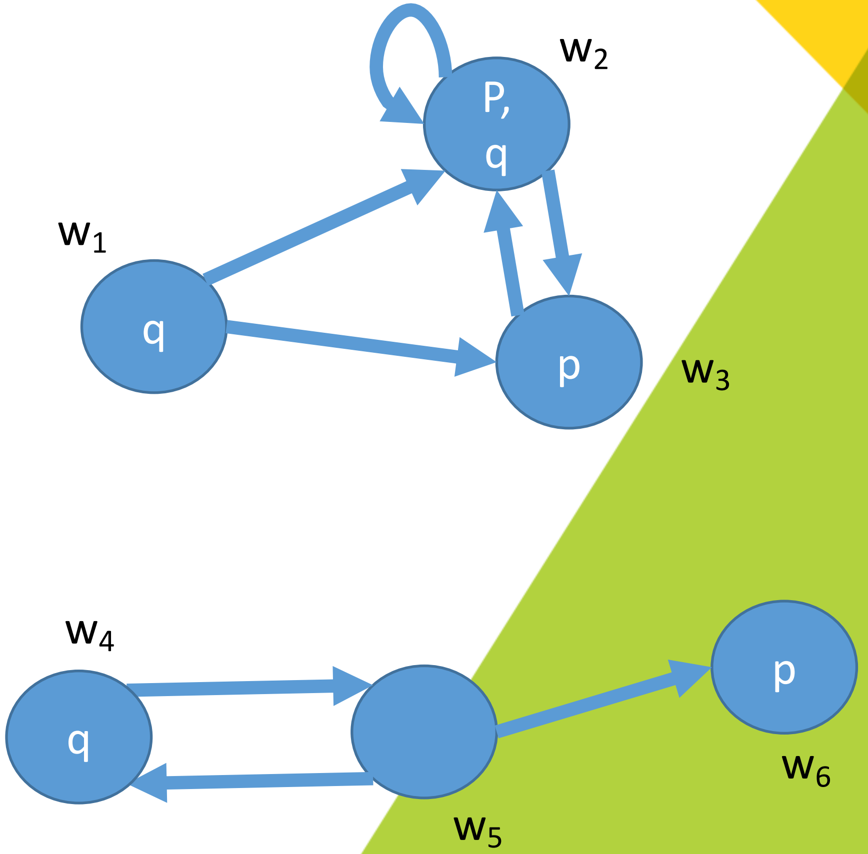
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\, \varphi$ iff $\forall\, w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall\, i \in G\ (w \Vdash K_i\, \psi)$

$w \Vdash C_G\varphi$ iff $\forall\, k \geq 1\ (w \Vdash E_G^k\, \psi)$

$w \Vdash D_G\varphi$ iff $\forall\, w' \in W\ (\forall\, i \in G\ (R_i\,(w, w') \rightarrow w \Vdash \psi))$



KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

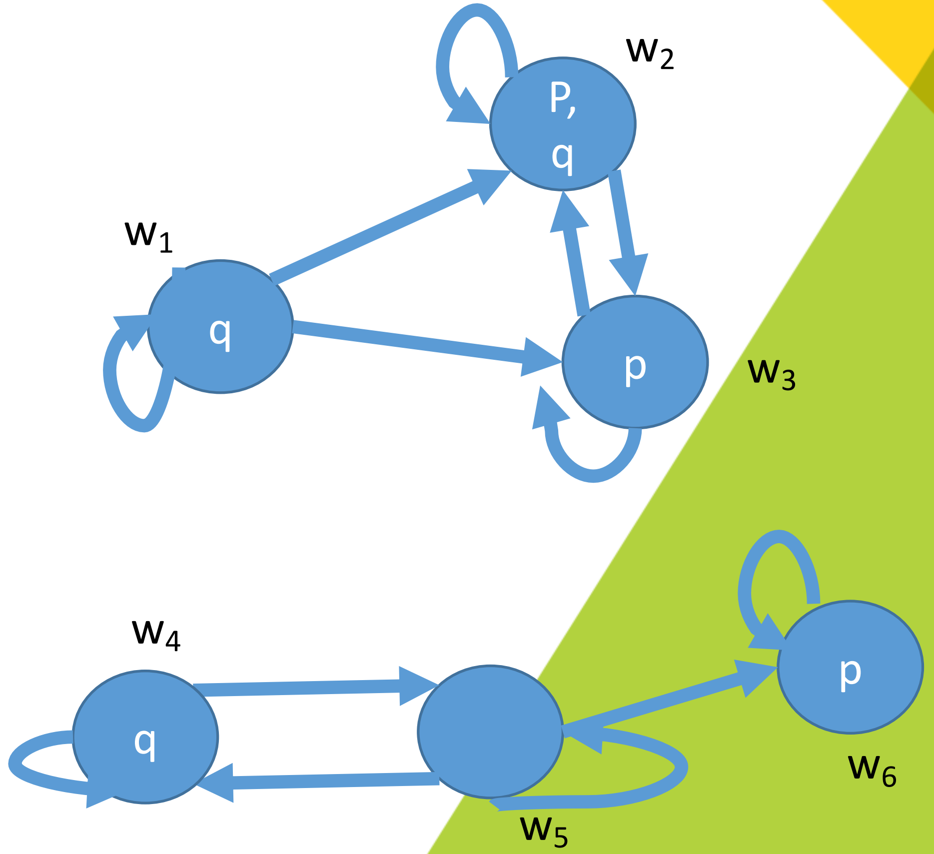
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\, \varphi$ iff $\forall w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G\ (w \Vdash K_i\, \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1\ (w \Vdash E_G^k\, \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W\ (\forall i \in G\ (R_i\ (w, w') \rightarrow w \Vdash \psi)$



$\Box\varphi \rightarrow \varphi$

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = \left(W, \{R\}_{\{i \in \mathrm{A}\}}, L\right)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

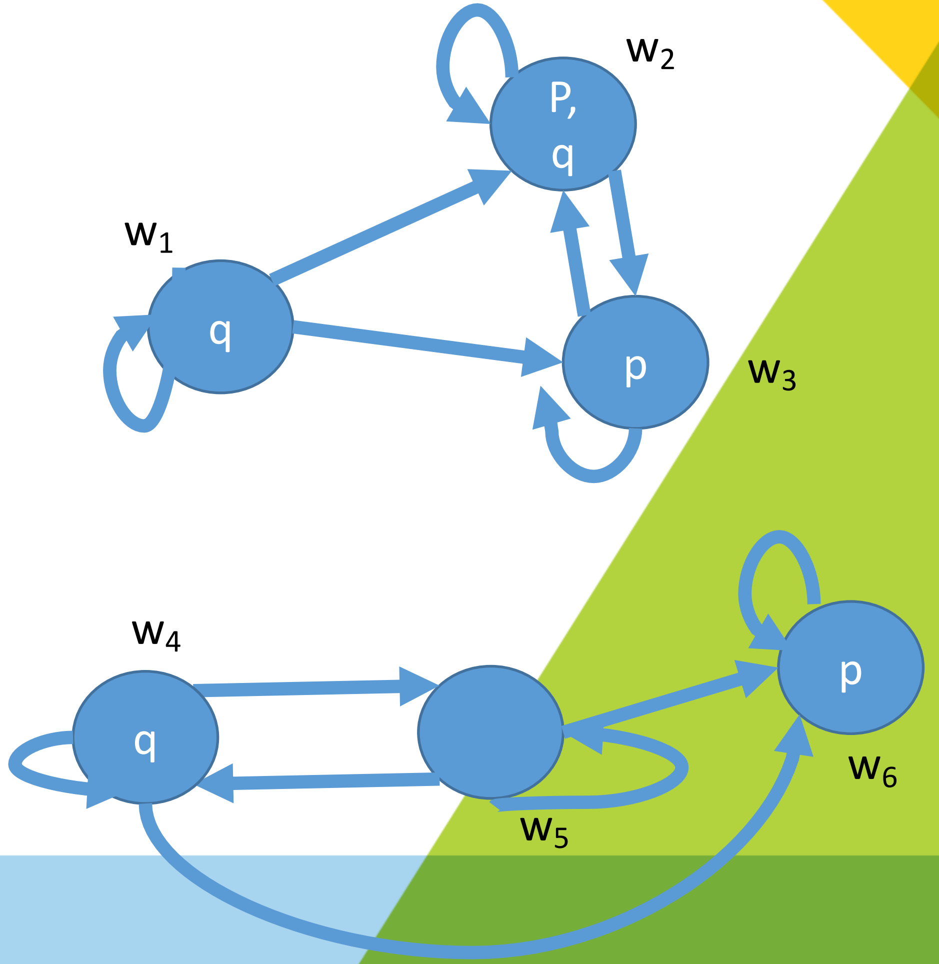
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\,\varphi$ iff $\forall w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G\ (w \Vdash K_i\,\psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1\ (w \Vdash E_G^k\,\psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W\ (\forall i \in G\ (R_i\ (w, w') \rightarrow w \Vdash \psi)$



$\Box\boldsymbol{\varphi} \rightarrow \Box\Box\boldsymbol{\varphi}$

KT45ⁿ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = \left(W, \{R\}_{\{i\in A\}}, L\right)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

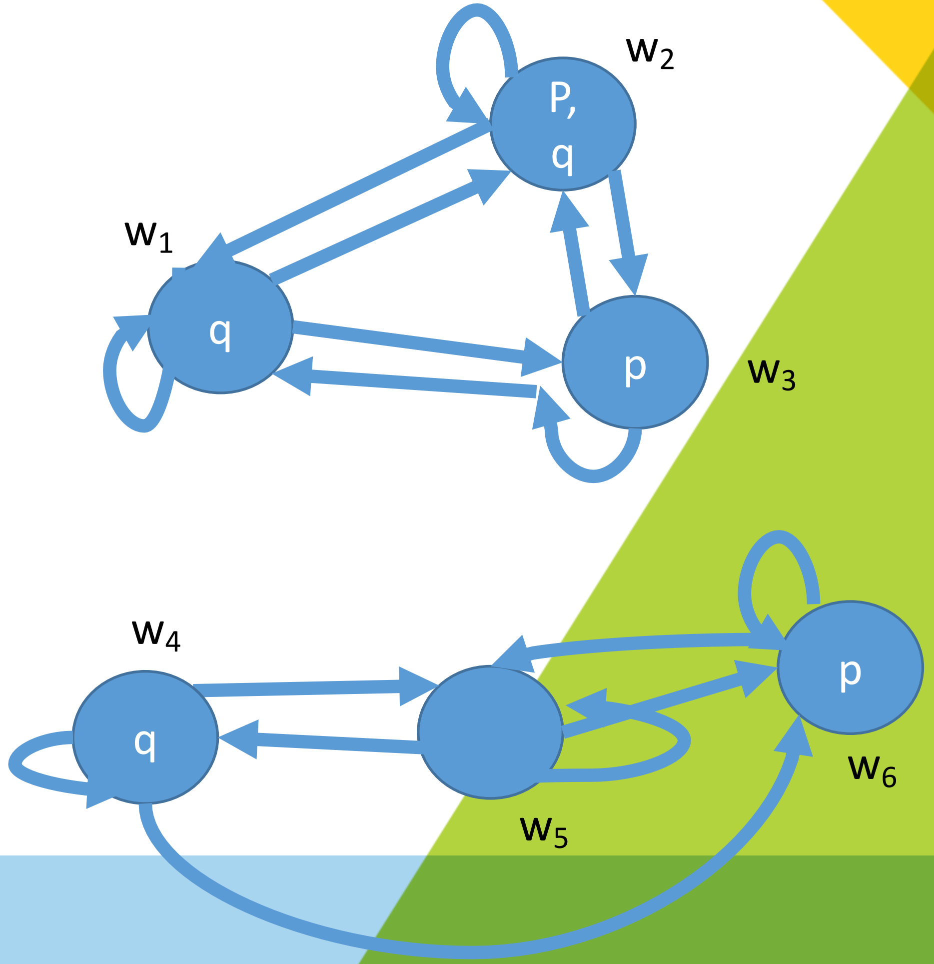
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\,\varphi$ iff $\forall w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G\ (w \Vdash K_i\,\psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1\ (w \Vdash E_G^k\,\psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W\ (\forall i \in G\ (R_i(w, w') \rightarrow w \Vdash \psi)$



$\diamond\boldsymbol{\varphi} \rightarrow \square\diamond\boldsymbol{\varphi}$

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i\in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

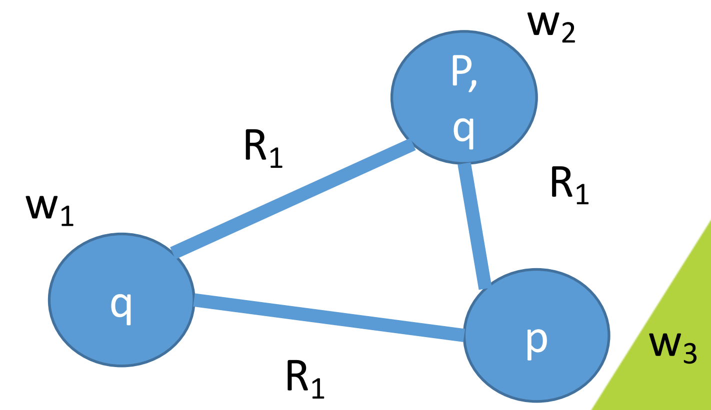
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

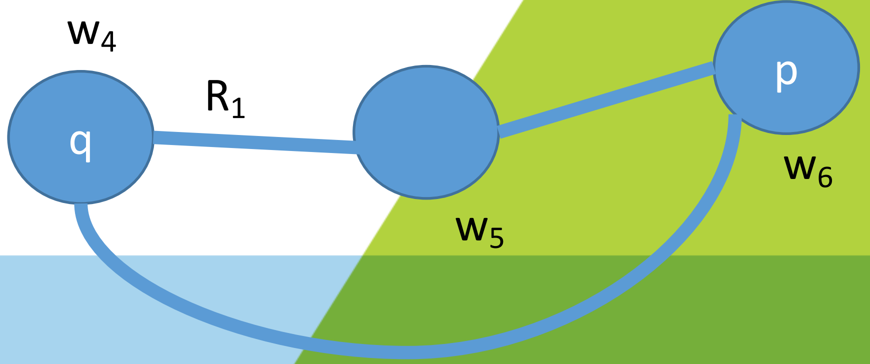
$w \Vdash K_i \varphi$ iff $\forall w' \in W \ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G \ (w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1 \ (w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W \ (\forall i \in G \ (R_i (w, w') \rightarrow w \Vdash \psi)$





For what world(s) w is $w \Vdash K_1 q$ true?

$KT45^n$ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

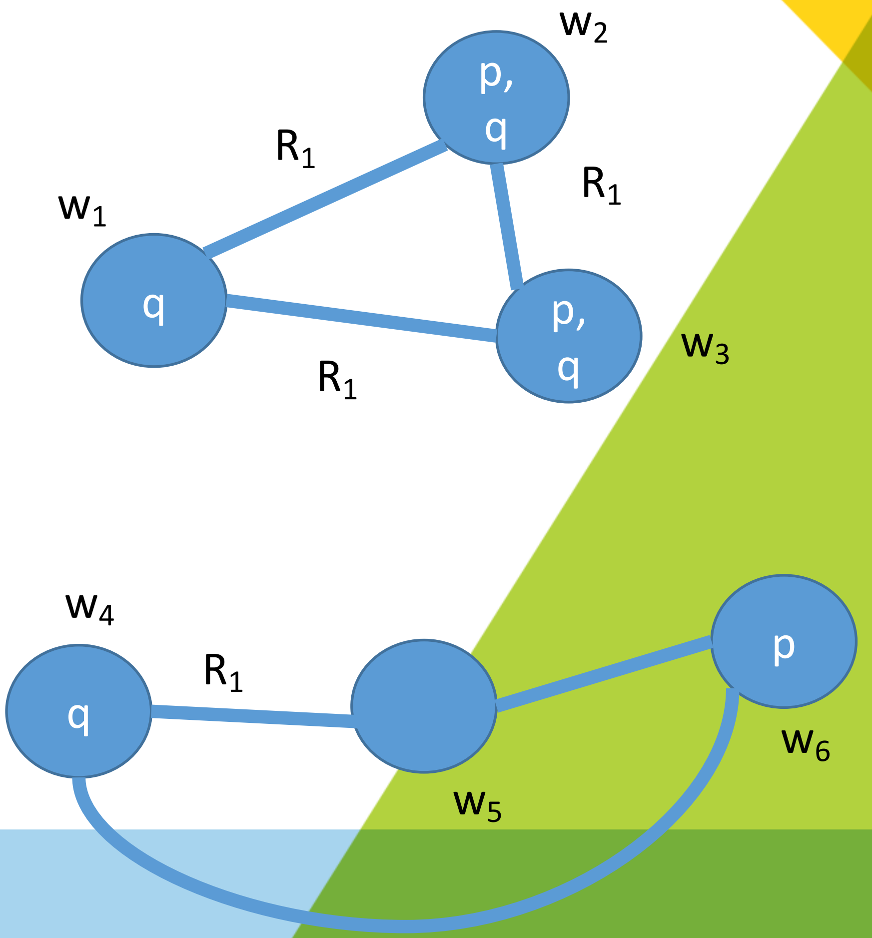
$w \Vdash \varphi \to \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\,\varphi$ iff $\forall\, w' \in W\ (R_i(w, w') \to w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall\, i \in G\ (w \Vdash K_i\,\psi)$

$w \Vdash C_G\varphi$ iff $\forall\, k \geq 1\ (w \Vdash E_G^k\,\psi)$

$w \Vdash D_G\varphi$ iff $\forall\, w' \in W\ (\forall\, i \in G\ (R_i\,(w, w') \to w \Vdash \psi)$



For what world(s) w is $w \Vdash K_1\, q$ true?

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

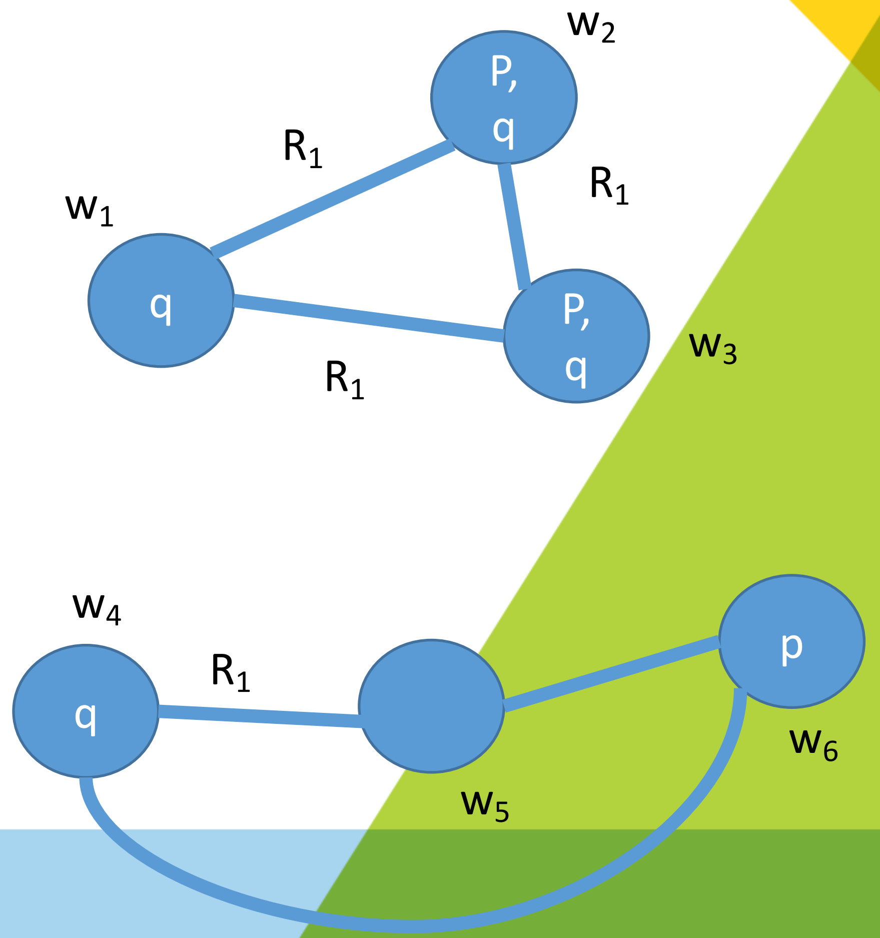
$w \Vdash \varphi \to \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i \varphi$ iff $\forall w' \in W$ $(R_i(w, w') \to w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G$ $(w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1$ $(w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W$ $(\forall i \in G$ $(R_i (w, w') \to w \Vdash \psi)$



$w \Vdash E_G$ q trivially true for G = {}

$\mathbf{KT45}^n$ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in \mathrm{A}\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

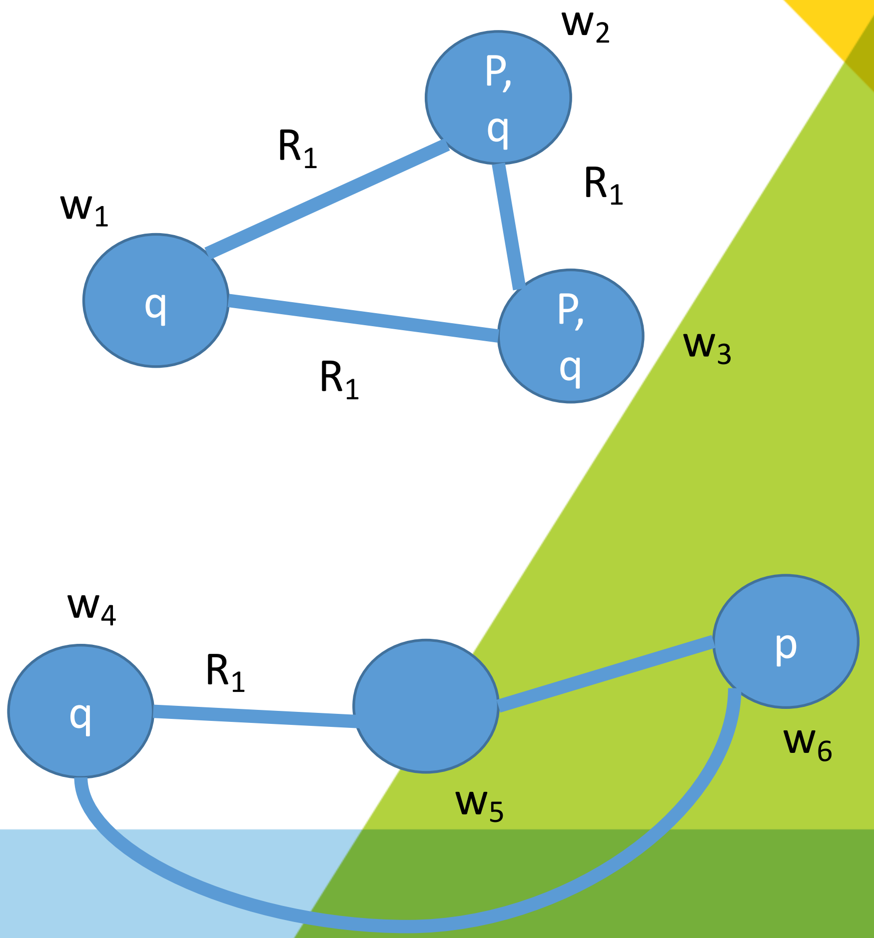
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i \varphi$ iff $\forall w' \in W$ $(R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G$ $(w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1$ $(w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W$ $(\forall i \in G$ $(R_i (w, w') \rightarrow w \Vdash \psi)$



$w \Vdash E_G$ q true for G = {1}

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

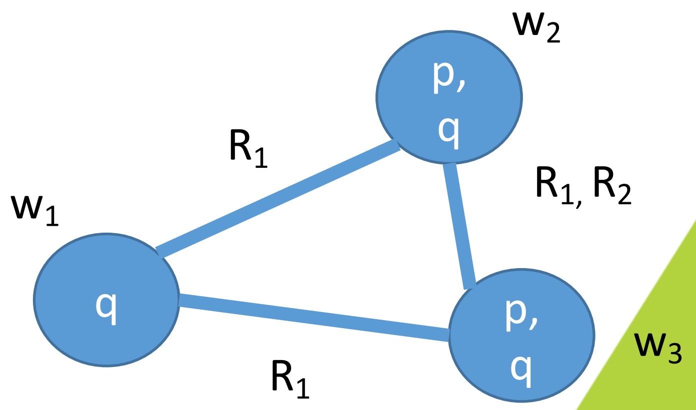
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

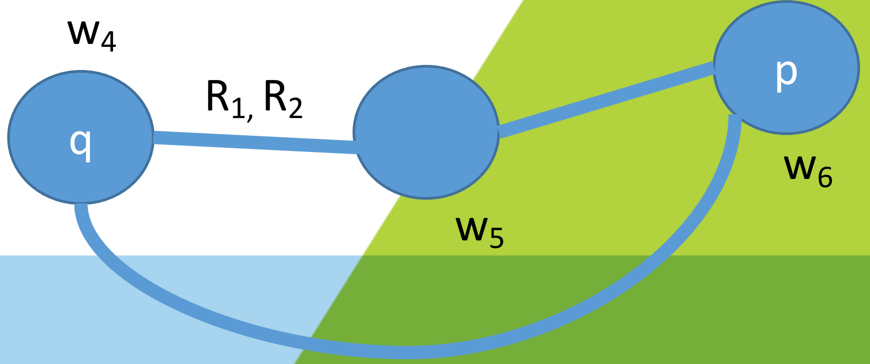
$w \Vdash K_i\,\varphi$ iff $\forall\, w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall\, i \in G\ (w \Vdash K_i\,\psi)$

$w \Vdash C_G\varphi$ iff $\forall\, k \geq 1\ (w \Vdash E_G^k\,\psi)$

$w \Vdash D_G\varphi$ iff $\forall\, w' \in W\ (\forall\, i \in G\ (R_i\,(w, w') \rightarrow w \Vdash \psi)$





For what worlds w is $w \Vdash E_G\, q$ true for $G = \{1, 2\}$?

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

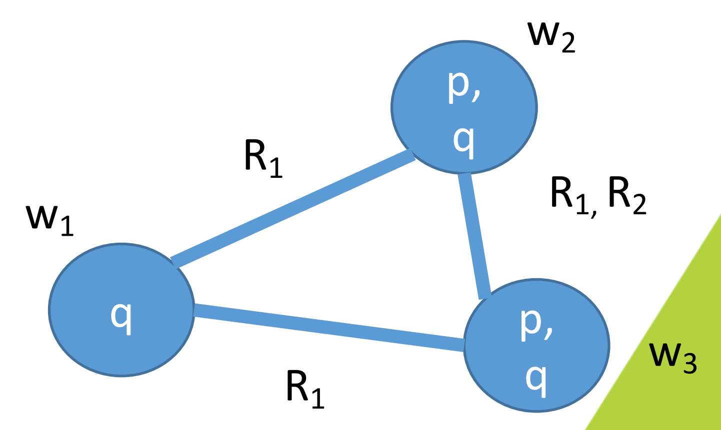
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

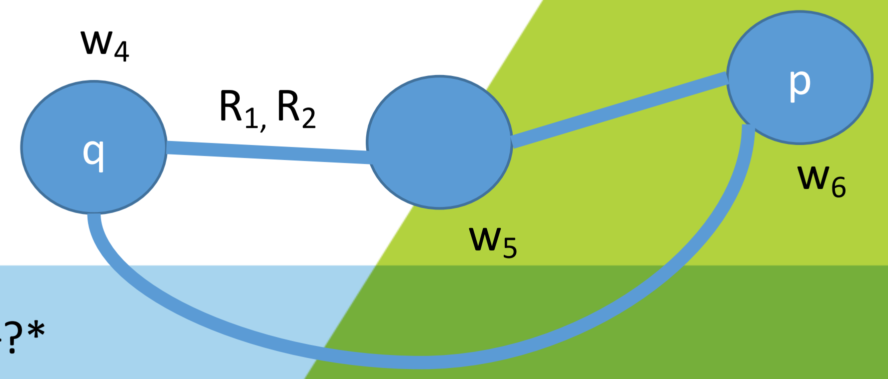
$w \Vdash K_i \varphi$ iff $\forall w' \in W\ (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G\ (w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1\ (w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W\ (\forall i \in G\ (R_i (w, w') \rightarrow w \Vdash \psi)$





For what worlds w is $w \Vdash E_G E_G\, q$ true for G = {1, 2}?*

KT45^n syntax & Semantics

$$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$$
$$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in \mathrm{A}\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

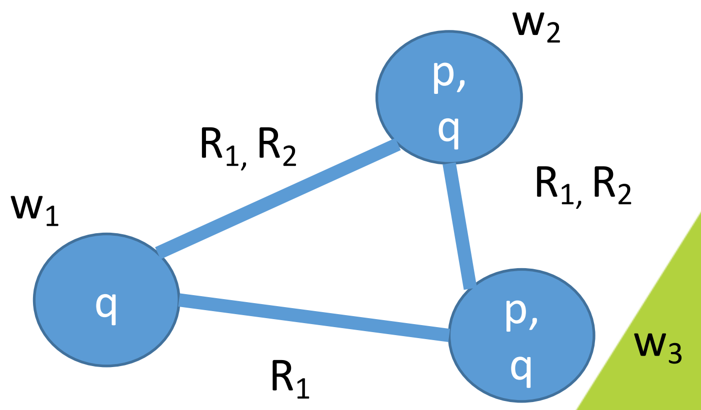
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

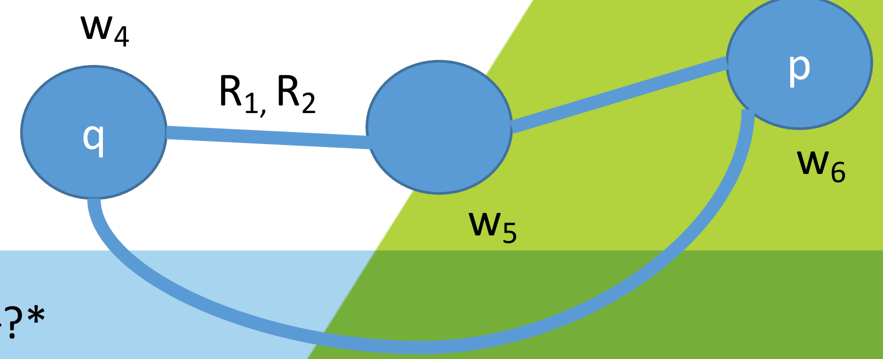
$w \Vdash K_i\, \varphi$ iff $\forall\, w' \in W\, (R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall\, i \in G\, (w \Vdash K_i\, \psi)$

$w \Vdash C_G\varphi$ iff $\forall\, k \geq 1\, (w \Vdash E_G^k\, \psi)$

$w \Vdash D_G\varphi$ iff $\forall\, w' \in W\, (\forall\, i \in G\, (R_i\, (w, w') \rightarrow w \Vdash \psi)$





For what worlds w is $w \Vdash E_G\, E_G\, q$ true for G = {1, 2}?*

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

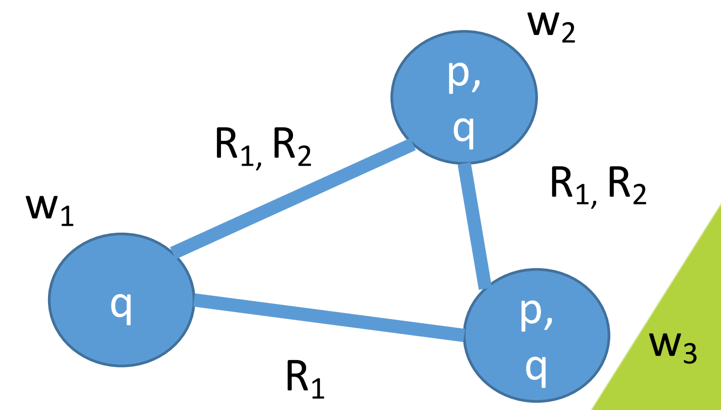
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

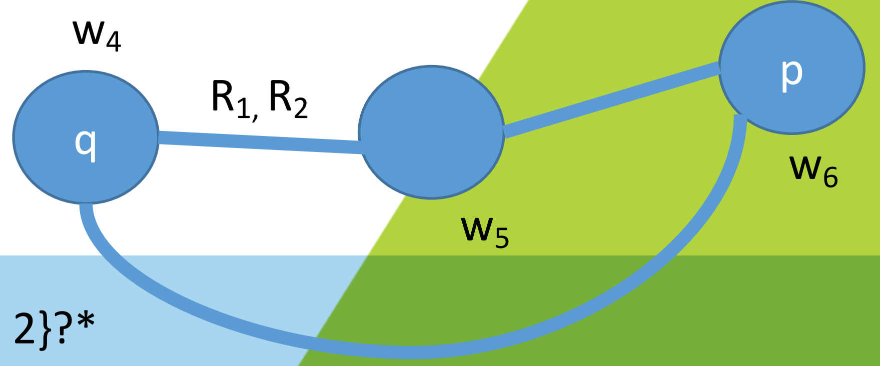
$w \Vdash K_i \varphi$ iff $\forall w' \in W$ $(R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G$ $(w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1$ $(w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W$ $(\forall i \in G$ $(R_i(w, w') \rightarrow w \Vdash \psi)$





For what worlds w is $w \Vdash E_G E_G E_G\, q$ true for $G = \{1, 2\}$?*

KT45^n syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2$
$\mid \mathrm{K}_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in \mathrm{A}\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

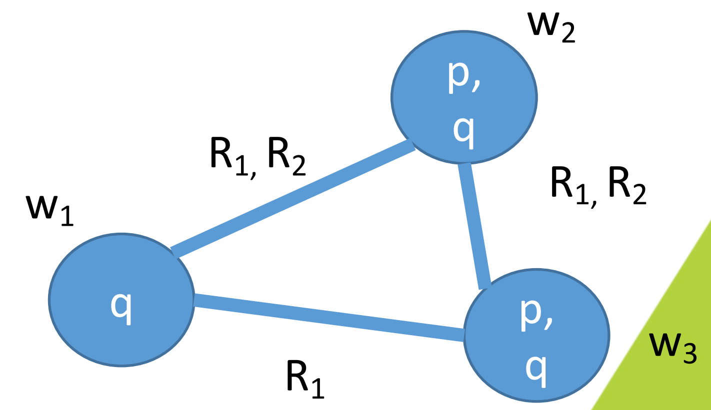
$w \Vdash \varphi \rightarrow \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

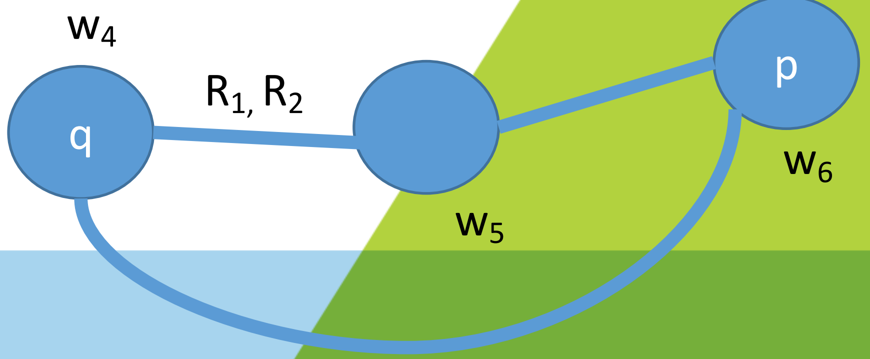
$w \Vdash K_i \varphi$ iff $\forall w' \in W$ $(R_i(w, w') \rightarrow w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall i \in G$ $(w \Vdash K_i \psi)$

$w \Vdash C_G\varphi$ iff $\forall k \geq 1$ $(w \Vdash E_G^k \psi)$

$w \Vdash D_G\varphi$ iff $\forall w' \in W$ $(\forall i \in G$ $(R_i(w, w') \rightarrow w \Vdash \psi)$





For what worlds w is $w \Vdash C_G\, q$ true for G = {1, 2}?

$KT45^n$ syntax & Semantics

$\top \mid \bot \mid p \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2$
$\mid K_i\varphi \mid E_G\varphi \mid C_G\varphi \mid D_G\varphi$

Given Model structure: $\mathcal{M} = (W, \{R\}_{\{i \in A\}}, L)$.

$w \Vdash \top$, $w \nVdash \bot$, $w \Vdash a$ iff $a \in L(w)$, $w \Vdash \neg\varphi$ iff $w \nVdash \varphi$

$w \Vdash \varphi \wedge \psi$ iff $w \Vdash \varphi$ and $w \Vdash \psi$

$w \Vdash \varphi \vee \psi$ iff at least one of $w \Vdash \varphi$ or $w \Vdash \psi$

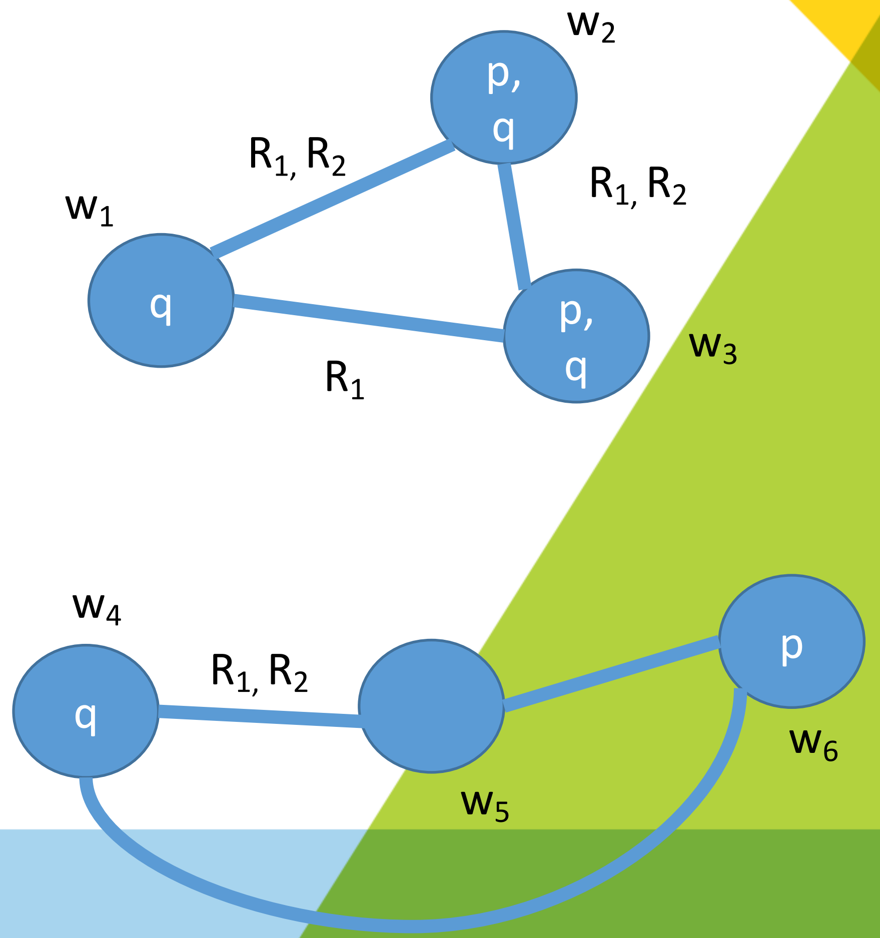
$w \Vdash \varphi \to \psi$ iff $w \Vdash \varphi$ whenever $w \Vdash \psi$

$w \Vdash K_i\,\varphi$ iff $\forall$ $w' \in W$ $(R_i(w, w') \to w' \Vdash \psi)$

$w \Vdash E_G\varphi$ iff $\forall$ $i \in G$ $(w \Vdash K_i\,\psi)$

$w \Vdash C_G\varphi$ iff $\forall$ $k \geq 1$ $(w \Vdash E_G^k\,\psi)$

$w \Vdash D_G\varphi$ iff $\forall$ $w' \in W$ $(\forall$ $i \in G$ $(R_i(w, w') \to w \Vdash \psi)$



For what worlds w is $w \Vdash D_G\, q$ true for G = {1, 2}?


The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

In Praise of Belief Bases: Doing Epistemic Logic without Possible Worlds

Emiliano Lorini
CNRS-IRIT, Toulouse University, France

Abstract

We introduce a new semantics for a logic of explicit and implicit beliefs based on the concept of multi-agent belief base. Differently from existing Kripke-style semantics for epistemic logic in which the notions of possible world and doxastic/epistemic alternative are primitive, in our semantics they are non-primitive but are defined from the concept of belief base. We provide a complete axiomatization and a decidability result for our logic.

Introduction

Epistemic logic and, more generally, formal epistemology are the areas at the intersection between philosophy (Hintikka 1962), artificial intelligence (AI) (Fagin et al. 1995; Meyer and van der Hoek 1995) and economics (Lismont and Mongin 1994) devoted to the formal representation of epistemic attitudes of agents including belief and knowledge. An important distinction in epistemic logic is between *explicit belief* and *implicit belief*. According to (Levesque 1984), "...a sentence is explicitly believed when it is actively held to be true by an agent and implicitly believed when it follows from what is believed" (p. 198). This distinction is particularly relevant for the design of resource-bounded agents who spend time to make inferences and do not believe all facts that are deducible from their actual beliefs.

The concept of explicit belief is tightly connected with the concept of *belief base* (Nebel 1992; Makinson 1985; Hansson 1993; Rott 1998). In particular, an agent's belief base, which is not necessarily closed under deduction, includes all facts that are explicitly believed by the agent. Nonetheless, existing logical formalizations of explicit and implicit beliefs (Levesque 1984; Fagin and Halpern 1987) do not clearly account for this connection.

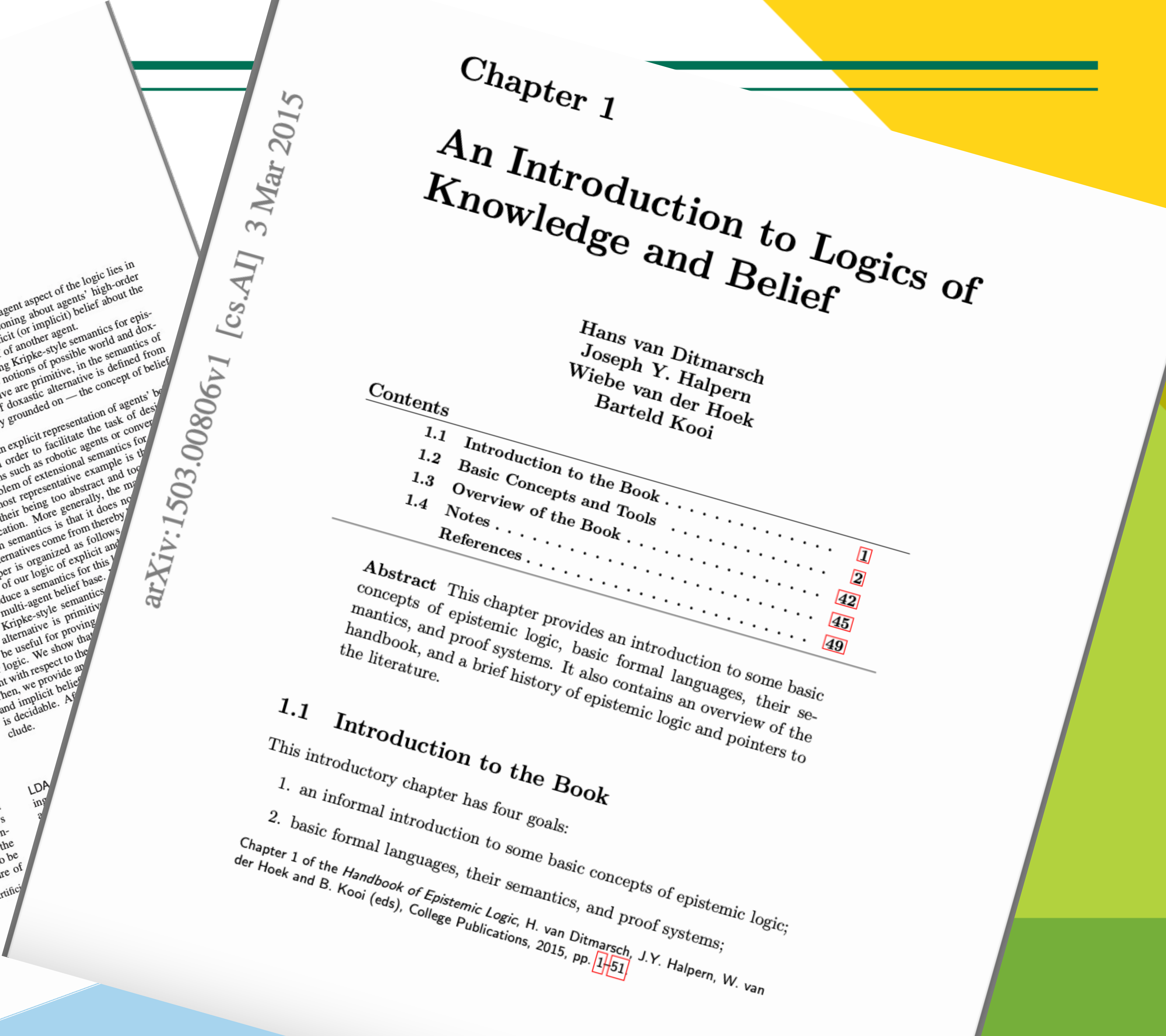
The aim of this paper is to fill this gap by providing a multi-agent logic that precisely articulates the distinction between explicit belief, as a fact in an agent's belief base, and implicit belief, as a fact that is deducible from the agent's explicit beliefs, given the agents' common ground. The concept of *common ground* (Stalnaker 2002) corresponds to the body of information that the agents commonly believe to be the case and that has has to be in the deductive closure of

their belief bases. The multi-agent aspect of the logic lies in the fact that it supports reasoning about agents' high-order beliefs, i.e., an agent's explicit (or implicit) belief about the explicit (or implicit) belief of another agent.

Differently from existing Kripke-style semantics for epistemic logic in which the notions of possible world and doxastic/epistemic alternative are primitive, in the semantics of our logic the notion of doxastic alternative is defined from — and more generally grounded on — the concept of belief base.

We believe that an explicit representation of agents' be
bases is crucial in order to facilitate the task of desi
intelligent systems such as robotic agents or conver
agents. The problem of extensional semantics for
logic, whose most representative example is th
semantics, is their being too abstract and too
agent specification. More generally, the ma
the Kripkean semantics is that it does no
doxastic alternatives come from thereby

The paper is organized as follows
language of our logic of explicit and
we introduce a semantics for this
tion of multi-agent belief base.
tional Kripke-style semantics
astic alternative is primitiv
will be useful for proving
our logic. We show that
lent with respect to the
Then, we provide an
and implicit belie
is decidable. A
clude.



Chapter 1

An Introduction to Logics of Knowledge and Belief

Hans van Ditmarsch
Joseph Y. Halpern
Wiebe van der Hoek
Barteld Kooi

Contents

Abstract This chapter provides an introduction to some basic concepts of epistemic logic, basic formal languages, their semantics, and proof systems. It also contains an overview of the handbook, and a brief history of epistemic logic and pointers to the literature.

1.1 Introduction to the Book

This introductory chapter has four goals:

1. an informal introduction to some basic concepts of epistemic logic;
2. basic formal languages, their semantics, and proof systems;

Chapter 1 of the *Handbook of Epistemic Logic*, H. van Ditmarsch, J.Y. Halpern, W. van der Hoek and B. Kooi (eds), College Publications, 2015, pp. 1–51.

Exercise: CGMs vs. Kripke structures

The University of Vermont