Modal logic

Michael Franke *October 2022*

Notions covered: ...

to be cleaned up and completed

1 Basics of Modal Logic

1.1 Language

Definition 1 (Syntax). Let $\Phi = \{p, q, r, \ldots\}$ be a set of proposition letters. The **language** $\mathfrak{L}_{\Phi,i}$ of basic modal logic based on Φ for i agents is:

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \wedge \psi \mid B_i\varphi \,.$$

Additionally, we allow the following notational variants:

$$\varphi \vee \psi \ \equiv \ \neg(\neg\varphi \wedge \neg\psi) \qquad \varphi \leftrightarrow \psi \ \equiv \ (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$$
$$\varphi \rightarrow \psi \ \equiv \ \neg\varphi \vee \psi \qquad P_i\varphi \ \equiv \ \neg B_i \neg\varphi \,.$$

1.2 Semantics

Definition 2 (Model). A **modal model** (aka. **Kripke structure**) for the language $\mathfrak{L}_{\Phi,i}$ is a triple $\mathcal{M} = \left\langle W, V, (R_j)_{j \leq i} \right\rangle$ such that:

- W is a set of **worlds**,
- $V : W \times \Phi \rightarrow \{0, 1\}$ is a **valuation function**,
- for each agent i $R_i \subseteq W \times W$ is an **accessibility relation**.

Notation and terminology:

- write wR_iv for $\langle w, v \rangle \in R_i$
- write $R_i(w) = \{v \in W \mid wR_iv\}$
- implicitly restrict attention to **serial models** where $R_i(w) \neq \emptyset$ for all w and i
- if $\mathcal{M} = \left\langle W, V, (R_j)_{j \leq i} \right\rangle$ and $w \in W$, we call the pair $\mathcal{M}, w$ a **pointed model**
- we say that a model is reflexive (transitive, Euclidean) if all of its relations R_i are reflexive (transitive, Euclidean)
- for $C \subseteq \{r, t, e\}$ let $\mathfrak{M}^C_{\Phi,i}$ be the class of all models $\mathcal{M}_{\Phi,i}$ that have the properties selected in C
  - $\mathfrak{M}^{\{t,e\}}_{\Phi,i}$ is the class of **belief models**

– $\mathfrak{M}^{\{r,t,e\}}_{\Phi,i}$ is the class of **knowledge models**

Definition 3 (Truth & Validity). Fix $\mathcal{M} = \left\langle W, V, (R_j)_{j\leq i}\right\rangle$ and $w \in W$ and define **truth in pointed models** inductively:

$$\mathcal{M}, w \models p \text{ iff } V(w, p) = 1 \qquad \mathcal{M}, w \models \varphi \wedge \psi \text{ iff } \mathcal{M}, w \models \varphi \text{ and } \mathcal{M}, w \models \psi$$
$$\mathcal{M}, w \models \neg\varphi \text{ iff } \mathcal{M}, w \not\models \varphi \qquad \mathcal{M}, w \models B_i\varphi \text{ iff } \mathcal{M}, v \models \varphi \text{ for all } v \in R_i(w)$$

We say that a formula φ is **valid in a model** $\mathcal{M}$, $\mathcal{M} \models \varphi$, if for all $w \in W$ of that model $\mathcal{M}, w \models \varphi$. We say that a formula φ is **valid in a class of models** $\mathcal{N}$, $\mathcal{N} \models \varphi$, if φ is valid on every model in $\mathcal{N}$.

Example 4 (Modal Model). We consider a simple modal model for $\Phi = \{p, q\}$ and two agents i and j. The following formulas are true at world w:

- $B_j(\neg p \wedge \neg q)$

  "j believes that p and q are false"

- $\neg p \wedge B_i p$

  "i falsely believes that p is true"

- $\neg B_i q \wedge \neg B_i \neg q$

  "i is uncertain about q"

- $\neg B_j p \wedge B_i B_j p$

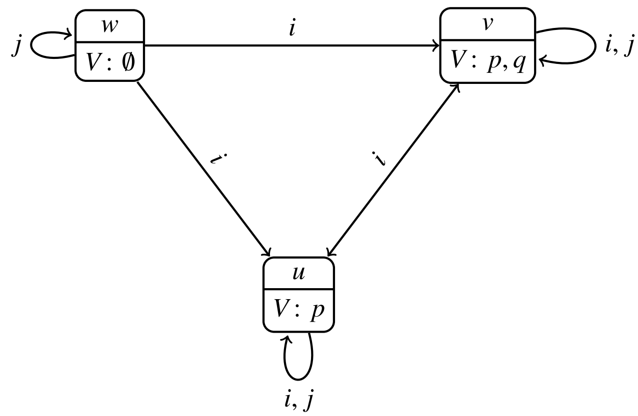
  "i falsely believes that j believes p"

- $\neg B_i B_j q \wedge \neg B_i \neg B_j q$

  "i is uncertain whether j believes q"

- $B_i(q \to B_j q \wedge \neg q \to B_j \neg q)$

  "i believes that j knows whether q"



1.3 Proof System

Definition 5 (Proof System for Modal Logics). A proof system for a modal logical language is given by a set of I **inference rules** and a set A of **axiom schemas**. Let I contain the rules:

MP: from φ and $\varphi \to \psi$ infer ψ (**modus ponens**)

Nec: from φ infer $B_i\varphi$ (**necessitation**) and let A always contain the **normal axiom schemata**:

Prop: all substitution instances of tautologies of propositional logic

K: $(B_i\varphi \wedge B_i(\varphi \rightarrow \psi)) \rightarrow B_i\psi$

and additionally any (possibly empty) subset of the following **additional axiom schemata**:

T: $B_i\varphi \rightarrow \varphi$ (**truth**)

4: $B_i\varphi \rightarrow B_iB_i\varphi$ (**positive introspection**)

5: $\neg B_i\varphi \rightarrow B_i\neg B_i\varphi$ (**negative introspection**)

Definition 6 (Proof System, Proof). A **proof** is a sequence of formulas each of which is either an axiom or derived from previous formulas in the sequence by application of an inference rule. For any (possibly empty) subset D of the additional axioms $\{\mathbf{T}, \mathbf{4}, \mathbf{5}\}$ we write $D \vdash \varphi$ if a proof exists for φ in the corresponding proof system.

Soundness & Completeness

Theorem 7 (Soundness & Completeness). Let C be a (possibly empty) subset of $\{r, t, e\}$ and let D the 'corresponding' subset of $\{\mathbf{T}, \mathbf{4}, \mathbf{5}\}$. Then:

$$\mathfrak{M}^C_{\Phi,i} \models \varphi \text{ iff } D \vdash \varphi\,.$$

The "left-to-right" part is **completeness**, the "right-to-left" part is **soundness**.