Logic and Computation II

Part 4. Modal logic

Kazuyuki Tanaka

BIMSA

March 25, 2025

北京雁栖湖
应用数学研究院
Beijing Institute
of Mathematical
Sciences and Applications

Logic and Computation II

- **Part 4. Modal logic**
- **Part 5. Modal μ-calculus**
- **Part 6. Automata on infinite objects**
- **Part 7. Recursion-theoretic hierarchies**

Part 4. Schedule (tentative)

- March 4, (1) Kripke models and normal logics
- March 6, (2) Kripke completeness
- March 11, (3) Standard translation and bisimulation
- March 13, (4) Decision problems
- March 18, (5) Computational complexity
- March 20, (6) Multi-modal and predicate logics
- March 25, (7) Modal predicate logics and epistemic logic (introduction to modal μ-calculus)

§4.8. Modal predicate logic

We define the formulas of modal predicate logic as follows:

$$\varphi ::= P(\vec{x}) \mid x = y \mid \neg\varphi \mid (\varphi \to \varphi) \mid \Box\varphi \mid \forall x\varphi.$$

- For simplicity, function symbols are omitted.
- Atomic formulas are n-ary predicates $P(\vec{x})$ and equality $x = y$.
- Other logical symbols $(\wedge, \vee, \Diamond, \exists)$ are defined as usual.

Definition 4.42 (Varying Domain Model)

A varying domain model $M = (W, R, D, V)$ consists of:

- (W, R): a frame as in modal propositional logic,
- D: a function assigning a domain D_s to each state s,
- V: an interpretation of relation symbol $P(\vec{v})$ as a subset $V(P, s) \subset D_s^n$, where n is the arity of $P(\vec{v})$.

To interpret a formula in a varying domain model, we use an assignment function g that maps each variable x to an element a in $\bigcup_s D_s$. More generally, g could be defined as a function mapping a pair (x, s) to an element a of D_s. But in the following, we assume that an individual a assigned to a variable x remains the same although a may appear or disappear across states. We denote by $g' \sim_x g$ that two assignments g and g' agree on all variables except x.

Definition 4.43

The satisfaction relation $M, g, s \models \varphi$ is defined as follows:

1. $M, g, s \models P(\vec{x}) \Leftrightarrow g(\vec{x}) \in V(P, s)$,
2. $M, g, s \models x = y \Leftrightarrow g(x) = g(y)$,
3. $M, g, s \models \neg\varphi$ and $(\varphi \to \varphi')$ are defined in the same way as propositional logic,
4. $M, g, s \models \Box\varphi \Leftrightarrow$ for all $t \in sR$ such that $g(\vec{x}) \in D_t^n$, $M, g, t \models \varphi$,
5. $M, g, s \models \forall x\varphi(x) \Leftrightarrow$ for all $g' \sim_x g$ such that $g'(x) \in D_s$, we have $M, g', s \models \varphi(x)$.

A formula $\varphi(\vec{x})$ is **valid in** M, denote $M \models \varphi(\vec{x})$, if for all $s \in W$ and all assignments g with $g(\vec{x}) \in D_s$, we have $M, g, s \models \varphi(\vec{x})$. If $\varphi(\vec{x})$ is valid in every model M, we say that $\varphi(\vec{x})$ is **valid**, and write $\models \varphi(\vec{x})$.

An important aspect of modal predicate logic is the validity of the following three formulas: **Barcan Formula** (BF), **Converse Barcan Formula** (CBF), and **Buridan Formula** (BuF).

$$\text{BF: } \forall x\Box\varphi \to \Box\forall x\varphi, \quad \text{CBF: } \Box\forall x\varphi \to \forall x\Box\varphi, \quad \text{BuF: } \exists x\Box\varphi \to \Box\exists x\varphi.$$

Lemma 4.44

A model M satisfies BF, i.e., $M \models$ BF, if and only if M has a decreasing domain property:

$$sRt \to D_t \subseteq D_s.$$

Proof. $(\Leftarrow)$ Assume that M has a decreasing domain and that $M, g, s \models \forall x\Box\varphi$.
So, for any $g' \sim_x g$ such that $g'(x) \in D_s$, we have $M, g', s \models \Box\varphi \ldots (i)$.
Then, for any $t \in sR$ such that $g'(x) \in D_t$, we have $M, g', t \models \varphi \ldots (ii)$.

Now to show $M, g, s \models \Box\forall x\varphi$, we first take any $t \in sR$ such that $g(x) \in D_t$, and then to prove $M, g, t \models \forall x\varphi$, we take any $g' \sim_x g$ such that $g'(x) \in D_t$. Since M has a decreasing domain, $g'(x) \in D_t$ implies $g'(x) \in D_s$, so $M, g', s \models \Box\varphi$ by (i), also $M, g', t \models \varphi$ by (ii). Since we took an arbitrary $g' \sim_x g$ such that $g'(x) \in D_t$, we have $M, g, t \models \forall x\varphi$. Since we took an arbitrary $t \in sR$ s.t. $g(x) \in D_t$, we finally have $M, g, s \models \Box\forall x\varphi$.

($\Rightarrow$) Suppose M does not have a decreasing domain. Then choose s, t such that sRt and a $d \in D_t - D_s$. Obviously, $s \neq t$. Define the interpretation of a unary predicate P as follows:

$$V(P,u) = D_u \text{ (for } u \neq t\text{)}, \quad V(P,t) = D_t - \{d\}$$

Then, $M, g, s \models \forall x \Box P$, but $M, g, t \not\models \forall x P$, contradicting BF. Thus, $M \not\models$ BF. □

By the definition of satisfaction (Definition 4.43, condition (5)), CBF holds in all models. Moreover, $M \models$ BuF holds if and only if M has an increasing domain property:

$$sRt \to D_s \subseteq D_t.$$

Thus, the validity of both BF and BuF implies that M has a constant domain, i.e.,

$$D_s \equiv \mathcal{D}, \quad \text{for all } s \in W.$$

The converse of the Buridan formula imposes a certain higher-order constraint on M.

- Condition (4) involves an interaction between the modal relation R and the fist-order domain D, making axiomatization complex.
- So, we often assume that M has an increasing domain, in which case condition (4) simplifies to:

  (4)′ $M, g, s \models \Box\varphi$ if and only if for all $t \in sR$, we have $M, g, t \models \varphi$.
- Under this assumption, the normal modal predicate logics K*, T*, S4*, etc., can be defined as natural extensions of propositional modal logics, and characterized in terms of frame properties.

Translation to Two-Sorted First-Order Logic

- In the case of constant domains, a formula φ of modal predicate logic can be translated into a formula $ST_x(\varphi)$ of two-sorted first-order logic as follows:

$$\begin{aligned} ST_x(P(\vec{v})) &:= P'(x, \vec{v}), \\ ST_x(u = v) &:= u = v, \\ ST_x(\neg\varphi) &:= \neg ST_x(\varphi), \\ ST_x(\varphi \to \psi) &:= ST_x(\varphi) \to ST_x(\psi), \\ ST_x(\Box\varphi) &:= \forall y(R(x, y) \to ST_y(\varphi)), \\ ST_x(\forall u\varphi) &:= \forall u ST_x(\varphi). \end{aligned}$$

  Here, a relational predicate $P'(s, \vec{v})$ is interpreted to mean that $\vec{v} \in V(P, s)$.

- A constant-domain model of modal predicate logic $M = (W, R, \mathcal{D}, V)$ can be associated with a first-order structure $M' = (W \uplus \mathcal{D}, R, P')$ (where $\uplus$ denotes the disjoint union).

Analogous to modal propositional logic, we have the following equivalences:

1. $M, s \models \varphi \Leftrightarrow M' \models ST_s(\varphi)$,
2. $M \models \varphi \Leftrightarrow M' \models \forall x ST_x(\varphi)$.

Thus, for constant-domain modal predicate logic, most properties of first-order structures remain applicable.

Note. While Buridan was a medieval thinker, but Barcan was a female logician who lived until about a decade ago. She was the first to attempt a predicate-logical formulation of Lewis's systems S2 and S4.

§4.9. Epistemic Logic EL

- Natural language contains many epistemic expressions such as “knows,” “believes,” and “doubts.”
- Logical research on these expressions has a long history, but the modern epistemic logic EL, based on Kripke models, originates from Hintikka’s “Knowledge and Belief” (1962).
- Since then, various formalizations and analyses have been developed. In the 1990s, Fagin and other theoretical computer scientists introduced *dynamic* epistemic logic, and in the 21st century, its connections to game theory have become prominent.
- In this lecture, we consider modal logic containing the modal operators, K_a and B_a.
  $\mathsf{K}_a\varphi$ means “agent a knows φ.”
  $\mathsf{B}_a\varphi$ means “agent a believes φ.”

Definition 4.45 (EL Formulas)

Given a set of atomic propositions P and a set of modal indices (agents) A, the formulas of EL are defined as follows:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \to \varphi) \mid \mathsf{K}_a\varphi \mid \mathsf{B}_a\varphi \qquad (p \in P, a \in A).$$

Other logical connectives are defined in the usual way:

$$\top := p \to p, \quad \varphi \vee \psi := \neg\varphi \to \psi, \quad \varphi \wedge \psi := \neg(\varphi \to \neg\psi).$$

Furthermore, the dual operators $\langle \mathsf{K}_a \rangle$ and $\langle \mathsf{B}_a \rangle$ are defined as:

$$\langle \mathsf{K}_a \rangle\varphi := \neg\mathsf{K}_a\neg\varphi, \quad \langle \mathsf{B}_a \rangle\varphi := \neg\mathsf{B}_a\neg\varphi.$$

- In a relational model $M = (W, (R^{\mathsf{K}}_a)_{a\in A}, (R^{\mathsf{B}}_a)_{a\in A}, v)$,
  $sR^{\mathsf{K}}_a t$ means “state t is consistent with agent a’s knowledge at state s,” and
  $sR^{\mathsf{B}}_a t$ means “state t is consistent with a’s belief at state s.”
- Let K_a be a necessity operator described by R^{K}_a. If K_a represents agent a’s “knowledge,” then the natural constraints on R^{K}_a are reflexivity (axiom T) and Euclideanity (axiom 5), making K_a an S5 operator.
- Let B_a be a necessity operator descibed by R^{B}_a. If B_a represents agent a’s “belief,” then R^{B}_a satisfies seriality (axiom D), transitivity (axiom 4), and Euclideanity (axiom 5), making B_a an KD45 operator. The key difference is that knowledge is always true, while belief may not be.
- Additionally, the following axioms are often assumed:
  KB1: $\mathsf{K}_a\varphi \to \mathsf{B}_a\varphi$
  KB2: $\mathsf{B}_a\varphi \to \mathsf{K}_a\mathsf{B}_a\varphi$
- Beyond these, questions arise regarding what assumptions lead to which conclusions.

Problem 6

Assume that $\mathsf{B}_a\varphi \to \mathsf{B}_a\mathsf{K}_a\varphi$. Show that this implies $\mathsf{B}_a\varphi \leftrightarrow \mathsf{K}_a\varphi$.

Muddy Children Puzzle

For an application of epistemic logic, let's look at the following puzzle.

Muddy Children Puzzle

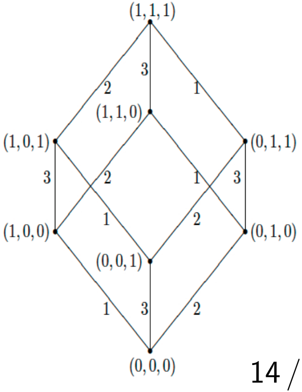
Three children have been playing outside. They can't see or feel if their own face is muddy, but they can see who of the others have mud on their face. As they come inside, mother tells them: 'Some of you have mud on your face. Do you know if you have mud on your face?" All three children say that they don't know. Mother asks again, "Do you know if you have mud on your face?" This time, two children say that they know. How many children have mud on their face? What happens if the mother asks her question three times?

- We formally express the above situation as follows. Let's name the three children $1, 2, 3$. Let an atomic proposition p_i represent "child i has mud on his face," and $\mathsf{K}_i\varphi$ represents "child i knows φ." We do not use the belief modality B_i here.
- For simplicity, we describe the state of the three children using a 3-tuple (b_1, b_2, b_3) where each b_i is either 0 or 1. For example, $(1, 0, 1)$ denotes the state where children 1 and 3 have mud while child 2 does not. That is, in the state $(1, 0, 1)$, p_1 and p_3 are true, whereas p_2 is false.

- Now, we analyze the knowledge of each child. In the state $(1,0,1)$, child 2 does not know whether he has mud. That is, child 2 cannot distinguish between states $(1,0,1)$ and $(1,1,1)$. This is represented as $(1,0,1)R_2^{\mathsf{K}}(1,1,1)$ or $(1,1,1)R_2^{\mathsf{K}}(1,0,1)$. Although $(1,0,1)R_2^{\mathsf{K}}(1,0,1)$ also holds, we omit mentioning such trivial cases, and describe this knowledge relation simply as:

$$(1,0,1) \underset{2}{\text{---}} (1,1,1)$$

- Using this, the knowledge relations over the eight possible states can be visualized as in the right:
- The relational model $M = (W, R_1^{\mathsf{K}}, R_2^{\mathsf{K}}, R_3^{\mathsf{K}}, v)$:
  - $W = \{(b_1, b_2, b_3) : b_i = 0, 1 (i = 1, 2, 3)\}$
  - $R_2^{\mathsf{K}} = \{((b_1, 0, b_3), (b_1, 1, b_3)) : b_1, b_3 = 0, 1\}$
  - $v((b_1, b_2, b_3), p_i) = \mathsf{T}$ (true) $\Leftrightarrow$ $b_i = 1$
  - $R_1^{\mathsf{K}}, R_3^{\mathsf{K}}$ are defined similarly.
- To find a general solution to this puzzle, we need to strengthen modal logic by introducing fixed-point operators. We will give a glimpse into this in the next slide, but we will present a complete solution after explaining modal logic with fixed-point operators, called **modal μ-calculus**, in the next few lectures.



- For a group of agents A, define:

$$\mathsf{E}\varphi := \bigwedge_{a\in A} \mathsf{K}_a\varphi$$

  which expresses "everyone knows φ". The operator E satisfies axioms (K, T) but does not necessarily satisfy (4, 5), as even if everyone happens to know a certain fact, it is possible that they do not know that others know it.
- Define iterated knowledge $\mathsf{E}^{n+1} = \mathsf{E}^n\mathsf{E}$ and introduce the operator C:

$$\mathsf{C}\varphi \equiv (\mathsf{E}\varphi \wedge \mathsf{E}^2\varphi \wedge \mathsf{E}^3\varphi \wedge \dots).$$

  A formula φ satisfying $\mathsf{C}\varphi$ is **common knowledge**, meaning not only everyone knows φ, but also everyone knows that everyone knows it, and so on. In "Muddy Children Puzzle," mere shared knowledge is insufficient; common knowledge is necessary.
- The operator C has a formal definition using the greatest fixed point operator ν:

$$\mathsf{C}\varphi \equiv \nu X.\mathsf{E}(\varphi \wedge X).$$

- Then, C satisfies the following properties:
  - Fixed-point axiom: $\mathsf{C}\varphi \to \mathsf{E}(\varphi \wedge \mathsf{C}\varphi)$.
  - Induction rule: If $\varphi \to \mathsf{E}(\varphi \wedge \psi)$ is a theorem, then $\varphi \to \mathsf{C}\psi$ is also a theorem.

Problem 7

Using the above axioms and rules, prove the following:
(1) $\mathrm{C}(\varphi \to \psi) \to (\mathrm{C}\varphi \to \mathrm{C}\psi)$.
(2) $\mathrm{C}(\varphi \to \mathrm{E}\varphi) \to (\varphi \to \mathrm{C}\psi)$.

Problem 8

While showing three prisoners some blue hats and red hats, the guard said: "I will now place one hat on each of your heads, but you will not be able to see the color of your own hat. There are only two red hats, so at least one of you will be wearing a blue hat. Now, if you are certain that your hat is blue, you may escape, but if it turns out to be red, you will be shot on the spot. What will you do?"

The three prisoners looked at each other for a while and then simultaneously ran away. Why were they all able to confidently determine that their own hats were blue? Consider this question based on the definition of common knowledge.

- Finally, we introduce "complete ignorance", the opposite of common knowledge. For simplicity, we consider the case of a single modal operator K.
- The operator $\mathsf{I}\varphi$, defined as follows, represents "**ignorance**" regarding φ:

$$\mathsf{I}\varphi \equiv \neg\mathsf{K}\varphi \wedge \neg\mathsf{K}\neg\varphi. \tag{1}$$

- Although I is not a normal modal operator, combinations such as KI and I^2 are intriguing. When using K in a negative way, it is more natural to consider a modal logic weaker than S5. However, if K satisfies S4, for any φ, $\neg\mathsf{I}^2\varphi$ becomes valid.
- In our paper[i], the following operation is introduced:

$$\alpha_\varphi(X) \equiv \neg\mathsf{K}(X \to \varphi) \wedge \neg\mathsf{K}(X \to \neg\varphi). \tag{2}$$

  Thus, $\alpha_\varphi := \alpha_\varphi(\top)$ coincides with $\mathsf{I}\varphi$ (where $\top = p \to p$). However, $\alpha_\varphi^2 := \alpha_\varphi(\alpha_\varphi)$ does not coincide with $\mathsf{I}^2\varphi$. Finally, we have the expression $\alpha_\varphi^\infty := \nu X.\alpha_\varphi(X)$, which can be considered to represent "**complete ignorance**."

---

[i] L.Pacheco and K.Tanaka, On the degrees of ignorance: via epistemic logic and μ-calculus, Proc. of SOCREAL 2022, 74-78.

Thank you for your attention!