Modal Logics II

Rodica Condurache

Lecture 2

Modal Logic - Recall

Definition (Basic Modal Logic - syntax)

Given a set $\mathcal{AP}$ of atomic propositions, an LP formula over $\mathcal{AP}$ is defined by the following syntax:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \Box\varphi \mid \Diamond\varphi$$

where $p \in \mathcal{AP}$.

If we want to keep the meaning open, we simply say "box" and "diamond". If we want to appeal to our intuition, we may say "necessarily" and "possibly" (or "forever in the future" and "sometime in the future")

Interpretation of Modalities

In a particular context $\Box\varphi$ could mean:

- It is necessarily true that φ;
- It will always be true that φ;
- It ought to be that φ;
- Agent Q believes that φ;
- Agent Q knows that φ;
- On any execution of program P, φ holds.

Since $\Box\varphi \equiv \neg\Diamond\neg\varphi$, we can infer the meaning of $\Diamond$ in each context.

Interpretation of Modalities

From the meaning of $\Box\varphi$, we can conclude the meaning of $\Diamond\varphi$, since $\Diamond\varphi \equiv \neg\Box\neg\varphi$:

$\Box\varphi$	$\Diamond\varphi$
It is necessarily true that φ	It is possibly true that φ
It will always be true that φ	Sometime in the future φ
It ought to be that φ	It is permitted to be that φ
Agent Q believes that φ	φ is consistent with Q's beliefs
Agent Q knows that φ	For all Q knows, φ
On any run of P, φ holds.	On some run of P, φ holds

Modalities lead to Interpretations of R

$\Box\varphi$	$R(x, y)$
It is necessarily true that φ	y is possible world according to info at x
It will always be true that φ	y is a future world of x
It ought to be that φ	y is an acceptable world according to the information at x
Agent Q believes that φ	y could be the actual world according to Q's beliefs at x
Agent Q knows that φ	y could be the actual world according to Q's knowledge at x
On any execution of P, φ holds	y is a possible resulting state after execution of P at x

Possible Properties of R

- **reflexive**: for every $w \in W$, we have R(x, x).
- **symmetric**: for every $x, y \in W$, we have R(x, y) implies R(y, x).
- **serial**: for every x there is a y such that R(x, y).
- **transitive**: for every $x, y, z \in W$, we have R(x, y) and R(y, z) imply R(x, z).
- **Euclidean**: for every $x, y, z \in W$ with R(x, y) and R(x, z), we have R(y, z).
- **functional**: for each x there is a unique y such that R(x, y).
- **linear**: for every $x, y, z \in W$ with R(x, y) and R(x, z), we have R(y, z) or y = z or R(z, y).
- **total**: for every $x, y \in W$, we have R(x, y) and R(y, x).
- **equivalence**: reflexive, symmetric and transitive.

Ignore L, and only consider the (W, R) part of a model, called frame.

Definition

A **frame** is a tuple $\mathcal{F} = \langle W, R \rangle$ where W is a set of worlds and $R \subseteq W \times W$ is a relation between worlds.

Definition

A Kripke structure $\mathcal{M}$ is said to be **defined over a frame** $\mathcal{F} = \langle W, R \rangle$ iff there is some labelling function $L : W \rightarrow 2^{\mathcal{AP}}$ such that $\mathcal{M} = \langle W, R, L \rangle$.

Definition

A basic modal logic formula φ is **valid in a frame** $\mathcal{F}$ ($\mathcal{F} \models \varphi$) iff φ is **globally true in any Kripke structure** $\mathcal{M}$ defined over $\mathcal{F}$.

Theorem

Let $\mathcal{F} = \langle W, R \rangle$ be a frame. The following statements are equivalent:

- *R is reflexive;*
- *$\Box\varphi \to \varphi$ is valid in $\mathcal{F}$, for any basic modal logic formula φ;*

Theorem

Let $F = (W, R)$ be a frame. The following statements are equivalent:

- *R is transitive;*
- *$\Box\varphi \to \Box\Box\varphi$ is valid in $\mathcal{F}$, for any basic modal logic formula φ;*

Formula Schemes and Properties of R

name	formula scheme	property of R
T	$\Box\varphi \rightarrow \varphi$	reflexive
B	$\varphi \rightarrow \Box\Diamond\varphi$	symmetric
D	$\Box\varphi \rightarrow \Diamond\varphi$	serial
4	$\Box\varphi \rightarrow \Box\Box\varphi$	transitive
5	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$	Euclidean
	$(\Box\varphi \rightarrow \Diamond\varphi) \wedge (\Diamond\varphi \rightarrow \Box\varphi)$	functional

Let $\mathcal{F}$ denote a set of formula scheme.

KT45 is characterized by : $\mathcal{F} = \{T, 4, 5\}$

name	formula scheme	property of R
T	$\Box\varphi \to \varphi$	reflexive
4	$\Box\varphi \to \Box\Box\varphi$	transitive
5	$\Diamond\varphi \to \Box\Diamond\varphi$	Euclidean

Used for reasoning about knowledge.

- T: **Truth**: agent Q only knows true things.
- 4: **Positive introspection**: If Q knows something, he knows that he knows it.
- 5: **Negative introspection**: If Q doesn't know something, he knows that he doesn't know it.

This is idealization of Knowledge ! Human knowledge has none of the properties above!!! Even computer agents can lack some of them.

Any sequence of modal operators and negations in KT45 is equivalent to one of the following: $-$, $\Box$, $\Diamond$, $\neg$, $\neg\Box$, *and* $\neg\Diamond$, *where* $-$ *indicates the absence of any negation or modality.*

Example

- $\Box\Box\varphi \equiv \Box\varphi$
- $\Diamond\Box\varphi \equiv \Diamond\varphi$
- $\neg\Diamond\neg\varphi \equiv \Diamond\varphi$

Multiagent systems

- KT45 can be applied to only one agent. That is usually not the case, in a multiagent system there are more then one agent.
- Reasoning in a multiagent system implies not only the knowledge an agent can extract directly from the system where he resides but also reasoning about what the other agents know.
- Ex. Muddy children puzzle, Prisoner dilemma etc.

Modal Logic $KT45^n$ - an epistemic modal logic for multiagent systems

- We consider now a set of agents: $A = \{1, 2, 3, ...n\}$
- Instead $\Box$, K_i will be used for expressing what i knows ($i \in \{1..n\}$).
- If we have $\mathcal{AP} = \{p, q, r, ..\}$, $K_i p$ means that the agent i knows that p is true.
- Ex: $K_1 p \wedge K_1 \neg K_2 K_1 p$ means: Agent 1 knows p, and also that Agent 2 does not know that Agent 1 knows p.

Syntax of Epistemic Modal Logic for multiagent systems

Definition

Let $A = \{1, 2, ..., n\}$ be a set of n agents. The formulas of **epistemic modal logic** are defined by

$$\varphi ::= \top \mid \bot \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid K_i\varphi$$

where $p \in \mathcal{AP}$ and $i \in A$.

Kripke structures for Epistemic Modal Logic in multiagent systems

Definition

A Kripke structure for the epistemic logic is a tuple $\mathcal{M} = \langle W, L, R_1, R_2, ...R_n \rangle$ where:

- W is a nonempty set of possible worlds;
- L is the labeling function that assigns to each state the set of propositions that are true

$$L : W \to 2^{\mathcal{AP}}$$

- R_i is a set of binary relations on W, $\forall i \in 1..n$ (a set of pairs, each pairs being formed from two elements of W).

Properties of accessibility relation

- A pair $(w_1, w_2) \in R_i$ means that the agent i cannot distinguish using his own sensors which of the world (w_1 or w_2) is the reality.

- Each relation between worlds is an equivalence relation:

  - reflexive (the agent cannot see differences in the same world)

  - symmetric (if the agent cannot distinguish between the worlds w_1 and w_2 then he also will not be able to distinguish between worlds w_2 and w_1)

  - transitive (if the agent cannot distinguish between worlds w_1 and w_2 and cannot distinguish the world w_2 from w_3, he will not be able to differentiate world w_1 from w_3)

Semantics of Epistemic Logic for multiagent systems

The semantics of a formula φ from the epistemic logic is defined with respect with the Kripke structure $\mathcal{M} = \langle W, L, R_1, R_2, ...R_n \rangle$ and the system's state $x \in W$, as it follows:

- $\mathcal{M}, x \models \top$
- $\mathcal{M}, x \not\models \bot$
- $\mathcal{M}, x \models p$ iff $p \in \pi(p)$
- ...
- $\mathcal{M}, x \models K_i\varphi$ iff for each $x' \in W$ with $(x, x') \in R_i$ we have that $\mathcal{K}, x \models \varphi$

Operators for groups of agents - $E_G\varphi$

- Some new connectors are found in Epistemic Logic for multiagent systems. Considering $G \subseteq A$ then we write:

- $E_G\varphi$ to symbolize that all agents in group G know that φ is true in the current system configuration.

- $E_G\varphi$ is read Everybody in the group G knows φ

- If $G = \{1, 2, 3, ..., m\}$ then

$$\mathcal{M}, x \models E_G\varphi \Leftrightarrow \mathcal{K}, x \models K_1\varphi \wedge K_2\varphi \wedge ...K_m\varphi$$

Operators for groups of agents - $C_G\varphi$

- If $G \subseteq A$ then:
- $C_G\varphi$ means that all agents in G know φ and each of them knows that all of them know it (this can be achieved only over a secure communication channels).
- $C_G\varphi$ is read Is common knowledge among agents in G that φ

$$\mathcal{M}, x \models C_G\varphi \Leftrightarrow \mathcal{M}, x \models E_G^k\varphi, \forall k \in N$$

$(E_G^0\varphi \equiv \varphi,\ E_G^1\varphi \equiv E_G\varphi,\ E_G^2\varphi \equiv E_G E_G\varphi,\ ...)$

Operators for groups of agents - $D_G\varphi$

If intelligent agents communicate with each other and use the knowledge each have, they can discover new knowledge.

- $D_G\varphi$ means that all agents in G can infer φ if they all put their knowledge in common.
- $D_G\varphi$ is read Is distributed knowledge among agents in G that φ

$$\mathcal{M}, x \models D_G\varphi \Leftrightarrow \mathcal{M}, y \models \varphi \text{ whenever } (x, y) \in R_i \text{ for all } i \in G.$$

$D_G\varphi$ - Example

- $\mathcal{K} = \langle W, \pi, K_1, K_2 \rangle$ where:
  - $W = \{w_1, w_2, w_3\}$;
  - $K_1 = \{(w_1, w_3)\}$, $K_2 = \{(w_1, w_2)\}$; (also $(k_1, k_1) \in K_1$. Why ?)
  - $\pi(w_1) = \{p, q, r\}$, $\pi(w_2) = \{p, q\}$, $\pi(w_3) = \{p, r\}$.

Which of the following is true: $(\mathcal{K}, w_1) \models K_1 p$, $(\mathcal{K}, w_1) \models K_1 q$, $(\mathcal{K}, w_1) \models K_1 r$, $(\mathcal{K}, w_1) \models K_2 p$, $(\mathcal{K}, w_1) \models K_2 q$, $(\mathcal{K}, w_1) \models K_2 r$, $(\mathcal{K}, w_1) \models D_{\{1,2\}} p$, $(\mathcal{K}, w_1) \models D_{\{1,2\}} q$, $(\mathcal{K}, w_1) \models D_{\{1,2\}} r$?

What can you say about $E_{\{1,2\}}$ or $C_{\{1,2\}}$ for the atomic propositions for w_1?

Muddy Children - 3 agents

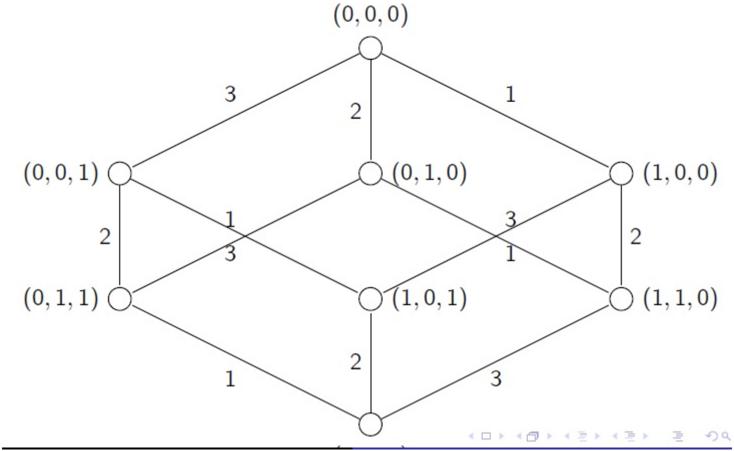
- $\mathcal{P} = \{p_1, p_2, p_3\}$ (Atomic propositions)
- $\mathcal{K} = \langle W, \pi, K_1, K_2, K_3 \rangle$;
- $W = \{(0,0,0), (0,0,1), ...(1,1,1)\}$; where $(0,0,1)$ means that only 3^{rd} children has mud on his forehead. $(|W| = 8)$
- $\pi((0,0,0)) = \emptyset$, $\pi((0,0,1)) = \{p_3\}$, $\ldots \pi((1,1,1)) = \{p_1, p_2, p_3\}$.

$K_1 = \{((0,0,0),(1,0,0)),\ ((0,1,0),(1,1,0)),\ ((0,0,1),(1,0,1)),\ ((0,1,1),(1,1,1))\}$

$K_2 = \{((0,0,0),(0,1,0)),\ ((0,0,1),(0,1,1)),\ ((1,0,0),(1,1,0)),\ ((1,0,1),(1,1,1))\}$

$K_3 = \{((0,0,0),(0,0,1)),\ ((1,0,0),(1,0,1)),\ ((0,1,0),(0,1,1)),\ ((1,1,0),(1,1,1))\}$

Muddy Children - 3 agents



Muddy Children - 3 agents

- The Kripke structure does not have arrows anymore (why?)
- There are no nodes with arcs to themselves (why?)
- Each edge will be labeled with the name of the agent considering the two states equivalent.

Muddy Children - 3 agents

Common Knowledge ($C_{\{1,2,3\}}$) - facts that are true in each state:

- $\models C_{\{1,2,3\}}(p_i \rightarrow K_j p_i)$, $\forall i, j$ with $i \neq j$;
- $\models C_{\{1,2,3\}}(\neg p_i \rightarrow K_j \neg p_i)$, $\forall i, j$ with $i \neq j$;

Muddy Children - 3 agents

$\varphi = p_1 \vee p_2 \vee p_3$ then $(K, (1, 0, 1)) \models E_{\{1,2,3\}}\varphi$?

$\varphi = p_1 \vee p_2 \vee p_3$ then $(K, (1, 0, 1)) \models E_{\{1,2,3\}}E_{\{1,2,3\}}\varphi$?

$\varphi = p_1 \vee p_2 \vee p_3$ then $(K, (1, 0, 1)) \models C_{\{1,2,3\}}\varphi$?

Muddy Children - 3 agents

After first round (father announces them that one is dirty):

That means that, in the current state of the system (let's call it w_{i1}), $(\mathcal{K}, w_{i1}) \models \varphi$ ($\varphi = p_1 \vee p_2 \vee p_3$). and allows to each agents to deduce that $(0, 0, 0)$ is not the current system's state ($C_{\{1,2,3\}}\varphi$).

In the Kripke structure, node representing the state $(0, 0, 0)$ and also edges to this state can be ignored / removed.

Muddy Children - 3 agents

After first round (father announces them that one is dirty):

That means that, in the current state of the system (let's call it w_{i1}), $(\mathcal{K}, w_{i1}) \models \varphi$ ($\varphi = p_1 \vee p_2 \vee p_3$). and allows to each agents to deduce that $(0,0,0)$ is not the current system's state ($C_{\{1,2,3\}}\varphi$).

In the Kripke structure, node representing the state $(0,0,0)$ and also edges to this state can be ignored / removed.

$$K_1 = \{((0,1,0),(1,1,0)),((0,0,1),(1,0,1)),((0,1,1),(1,1,1))\}$$

$$K_2 = \{((0,0,1),(0,1,1)),((1,0,0),(1,1,0)),((1,0,1),(1,1,1))\}$$

$$K_3 = \{((1,0,0),(1,0,1)),((0,1,0),(0,1,1)),((1,1,0),(1,1,1))\}$$

Muddy Children - 3 agents

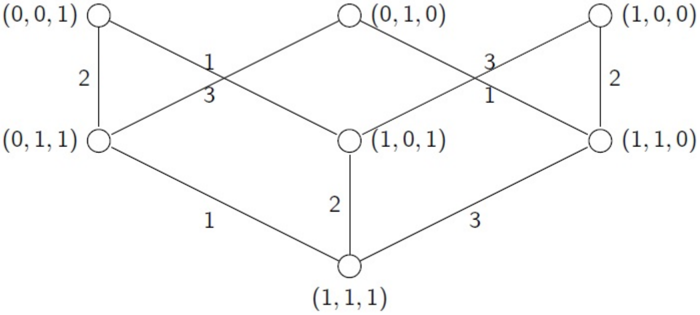


Figure: Kripke structure after 1^{st} step of muddy children puzzle

Muddy Children - 3 agents

If only one child would have been muddy (suppose is the 3^{rd} one), then the real system's state is $(0, 0, 1)$ and because child 3 no longer thinks the system state is $(0, 0, 0)$ or $(0, 0, 1)$ because $(0, 0, 0)$ was eliminated, he can successfully announce that he is dirty.

Supposing that there were two muddy children, none of the children would be able to correctly reason that he is dirty. After the second round, however they will understand that there are more then one muddy children (because otherwise the muddy one would have been capable to be sure that he is muddy).

So, if none of them deduce that he is muddy, in the second round, the worlds labeled with one atomic proposition are removed.

Muddy Children - 3 agents

$$K_1 = \{((0,1,1),(1,1,1))\}$$
$$K_2 = \{((1,0,1),(1,1,1))\}$$
$$K_3 = \{((1,1,0),(1,1,1))\}$$

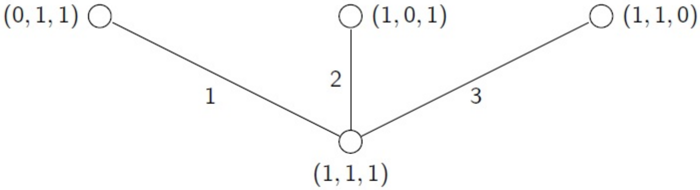


Figure: Kripke structure after 2^{nd} step of muddy children puzzle

Muddy Children - 3 agents

If the two that are dirty still aren't able to deduce that they are the dirty one, then, maybe, each of them expect the other two to come forward and since they don't do that, they all are simultaneously able (in the 3rd round) to tell that all of them are muddy.