

Modal logics: overview

- Part I: Introduction to modal and multimodal logics
 1. Motivation and introduction
 2. The basic multimodal logic K
 3. The basic monomodal logics
 4. Completeness of $G(k, l, m, n)$ logics, and decidability of the basic modal logics
 5. Basic multimodal logics
 6. Other modal logics
- Part II: Applications
 7. **Knowledge and announcements**
 8. Belief
 9. Common knowledge and common belief
 10. Action and propositional dynamic logic
 11. Goals and intentions
 12. Ability, agency and branching time
- Part II: Proof methods
 13. Translation method
 14. Tableau method

Part II: Applications

Outline of Part II: Applications

key concepts in agent interaction:

1 agent

⇒ mental state:

- ▶ knowledge and belief
- ▶ proactive attitudes: desire, goal, preference, intention
 - ★ BDI logics

2 action and time

- ▶ particular actions: speech acts

3 social level

- ▶ common knowledge, common belief
 - ★ reputation and trust
- ▶ joint goals, joint intentions
- ▶ norms: obligations, permissions, rights
- ▶ games, protocols

Chapter 7. Knowledge and announcements

Oct.14, 2008

Outline of chapter

- multiagent epistemic logic *KT5*
- public announcement logic *PAL*
 - ▶ the 'full' muddy children puzzle
- dynamic epistemic logics *DEL*

Epistemic logic: language

- *episteme* = ‘know’ (Greek)
- multiagent
- finite parameter set = set of agents
$$\begin{aligned}Prms &= \{i_1, i_2, \dots, i_{\text{card}(Agts)}\} \\ &= Agts\end{aligned}$$
(‘individuals’)
- knowledge explained in terms of possible worlds:
$$\begin{aligned}\Box_i \varphi &= \text{“agent } i \text{ knows that } \varphi\text{”} \\ &= \varphi \text{ true in every world that is possible for } i \\ &= K_i \varphi\end{aligned}$$
[Hintikka]
- 3 possible epistemic attitudes w.r.t. a formula φ :
 - ▶ $K_i \varphi$
 - ▶ $K_i \neg \varphi$
 - ▶ $\neg K_i \varphi \wedge \neg K_i \neg \varphi$

Knowledge: logical properties

standard logic of knowledge = multimodal $KT5$

- principles of multimodal K
- knowledge implies truth:
 - ▶ $K_i \varphi \rightarrow \varphi$ axiom T(K)
- positive introspection:
 - ▶ $K_i \varphi \rightarrow K_i K_i \varphi$ axiom 4(K)
- negative introspection:
 - ▶ $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$ axiom 5(K)

Knowledge: omniscience

knowledge set of agent i = set of formulas known by i

- i 's knowledge set is. . .

- ▶ closed under theorems:

- ★ $\frac{\varphi}{K_i \varphi}$

rule RN(K)

- ▶ closed under logical implication:

- ★ $\frac{\varphi \rightarrow \psi}{K_i \varphi \rightarrow K_i \psi}$

rule RM(K)

- ▶ closed under material implication:

- ★ $(K_i \varphi \wedge K_i (\varphi \rightarrow \psi)) \rightarrow K_i \psi$

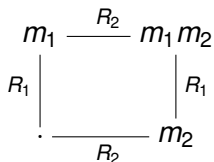
axiom K(K)

- *omniscience problem*

- ▶ if I know the axioms and inference rules of Peano Arithmetic then I know whether Goldbach's conjecture is true or not
- ▶ *KT5* is an idealization: rational agent, perfect reasoner
- ▶ inadequate for human agents
- ▶ widely accepted in AI

The logic of knowledge: semantics

- equivalence relation $R_i =$ indistinguishability for i
 $R_i(w)$ = “ i ’s alternatives to w ”
= “set of worlds i cannot distinguish from w ”
= “set of worlds compatible with i ’s knowledge”
= “*knowledge state* of agent i at w ”
- muddy children puzzle, initial situation (reflexive arrows omitted):



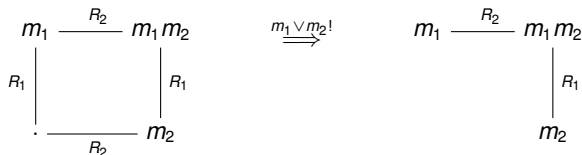
$$M, (m_1 m_2) \Vdash m_1 \wedge m_1 \wedge K_1 m_2 \wedge \neg K_1 m_1 \wedge \neg K_1 \neg m_1$$

The logic of knowledge: properties

- sound and complete: $\vdash_{KT5} \varphi$ iff $\models_{KT5} \varphi$
- decidable
- complexity of $KT5$ -satisfiability is
 - ▶ NP-complete if $\text{card}(Agts) = 1$
 - ▶ PSPACE-complete if $\text{card}(Agts) > 1$
- normal form only for monoagent $KT5$ (but not if $\text{card}(Agts) > 1!$)

Public announcement logic *PAL*

- observe: after the children have heard father's announcement that $m_1 \vee m_2$, they eliminate all those worlds where $m_1 \vee m_2$ is false
- idea: public announcements transform the model ('update')
- example of muddy children puzzle: father says " $m_1 \vee m_2$!"



Public announcement logic *PAL*: language

- parameter of modal operator = announcement of truth of $\varphi = \varphi!$
- roughly:

$$Prms = \{K_{i_1}, \dots, K_{i_{\text{card}(Agts)}}\} \cup \{\varphi! : \varphi \text{ is a formula}\}$$

- ▶ either circular definition of formulas
- ▶ or would not allow complex announcements
 - ★ $[(p \wedge \neg K_1 p)!]K_1 p$
 - ★ $[[p!]q!]K_i q$

- BNF:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i \varphi \mid [\varphi!]\varphi$$

where p ranges over *Atms* and i over *Agts*

- notations:

$$K_i \varphi = \Box_{K_i} \varphi$$

$$[\varphi!]\psi = \Box_{\varphi!} \psi$$

= “ ψ is true after every possible execution
of the announcement that φ is true”

= “after public announcement of φ , ψ is true”

$$\langle \varphi! \rangle \psi = \neg[\varphi!]\neg\psi$$

Public announcement logic *PAL*: models

- *PAL*-model = *KT5*-model
- truth conditions:
 - $M, w \Vdash p$ iff $w \in V(p)$
 - $M, w \Vdash \neg\varphi$ iff ...
 - $M, w \Vdash \varphi \wedge \psi$ iff ...
 - $M, w \Vdash K_i \varphi$ iff $M, w' \Vdash \varphi$ for all $w' \in R_i(w)$
 - $M, w \Vdash [\varphi!]\psi$ iff $M, w \not\Vdash \varphi$ or $M^{\varphi!}, w \Vdash \psi$
- “update of M by φ ” = $M^{\varphi!} = \langle W^{\varphi!}, R^{\varphi!}, V^{\varphi!} \rangle$, where
 - $W^{\varphi!} = \{w' \in W : M, w' \Vdash \varphi\}$
 - $R_i^{\varphi!} = R_i \cap (W^{\varphi!} \times W^{\varphi!})$
 - $V^{\varphi!}(p) = V(p) \cap W^{\varphi!}$
- *Remarks.*
 - ▶ announcements have to be truthful
 - ★ else satisfaction relation \Vdash would be ill-defined
 - ▶ if there is $w \in W$ such that $M, w \Vdash \varphi$ then $M^{\varphi!}$ is a *KT5*-model
- *PAL*-validity ($\Vdash_{PAL} \varphi$), *PAL*-satisfiability: as usual

Public announcements: validities

public announcements. . .

- are normal modal operators:
 - ▶ for all formulas φ, ψ , if $\models_{PAL} \psi \rightarrow \psi'$ then $\models_{PAL} [\varphi!]\psi \rightarrow [\varphi!]\psi'$
 - ▶ . . .
- are possible if true:
 $\models_{PAL} \varphi \rightarrow \langle \varphi! \rangle \top$
- are impossible if false:
 $\models_{PAL} \neg \varphi \rightarrow [\varphi!] \perp$
- are deterministic:
 $\models_{PAL} \langle \varphi! \rangle \psi \rightarrow [\varphi!] \psi$
- preserve the facts of the world:
if ψ is Boolean then $\models_{PAL} \psi \rightarrow [\varphi!] \psi$
- guarantee ‘no forgetting’ (alias ‘perfect recall’):
 $\models_{PAL} K_i [\varphi!] \psi \rightarrow [\varphi!] K_i \psi$
- guarantee ‘no learning’ (alias ‘no miracle’):
 $\models_{PAL} (\varphi \wedge [\varphi!] K_i \psi) \rightarrow K_i [\varphi!] \psi$

Public announcements: non-validities!

- public announcements do not always preserve knowledge:

$$\not\models_{PAL} K_i \psi \rightarrow [\varphi!] K_i \psi$$

- ▶ consider $\psi = \neg K_i p \dots$

- public announcements are not always successful:

$$\not\models_{PAL} [\varphi!] K_i \varphi$$

- ▶ consider $\varphi = p \wedge \neg K_i p$ ('Moore sentence'),
and remember: $K_i (p \wedge \neg K_i p)$ is *KT5*-unsatisfiable!

Axiomatics of *PAL*

- $KT5(K_i)$ = axiom schemas and inference rules of *KT5* for K_i : ...
- $K([\varphi!])$ = axiom schemas and inference rules of *K* for $[\varphi!]$: ...
- axiom schemas proper to announcements:
 - ▶ $\varphi \rightarrow \langle \varphi! \rangle \top$ (executability)
 - ▶ $\neg \varphi \rightarrow [\varphi!] \perp$ (inexecutability)
 - ▶ $\langle \varphi! \rangle \psi \rightarrow [\varphi!] \psi$ (determinism)
 - ▶ $\psi \rightarrow [\varphi!] \psi$ if ψ is atomic (pos. atomic preservation)
 - ▶ $\neg \psi \rightarrow [\varphi!] \neg \psi$ if ψ is atomic (neg. atomic preservation)
 - ▶ $K_i [\varphi!] \psi \rightarrow [\varphi!] K_i \psi$ (no forgetting)
 - ▶ $(\varphi \wedge [\varphi!] K_i \psi) \rightarrow K_i [\varphi!] \psi$ (no learning)
- *PAL*-theorem ($\vdash_{PAL} \varphi$), *PAL*-consistency: defined as usual
- soundness:
 - ▶ all axiom schema instances are *PAL*-valid
 - ▶ all inference rules preserve *PAL*-validity
- completeness: proof will use reduction axioms ...

Axiomatics of *PAL*: useful equivalences

- $\vdash_{PAL} [\varphi!] \psi \leftrightarrow (\neg \varphi \vee \psi)$ if ψ is Boolean

proof: induction on structure of ψ

for ψ atomic:

1 $\neg \psi \rightarrow [\varphi!] \neg \psi$

2 $([\varphi!] \psi \wedge \neg \psi) \rightarrow [\varphi!] \perp$

3 $[\varphi!] \perp \rightarrow \neg \varphi$

4 $[\varphi!] \psi \rightarrow (\neg \varphi \vee \psi)$

5 $\neg \varphi \rightarrow [\varphi!] \psi$

6 $\psi \rightarrow [\varphi!] \psi$

neg. atomic preservation

from 1. by logic $K([\varphi!])$

executability

from 2. and 3. by CPL

inexecutability

pos. atomic preservation

- $\vdash_{PAL} [\varphi!] \neg \psi \leftrightarrow (\neg \varphi \vee \neg [\varphi!] \psi)$

proof:

1 $([\varphi!] \neg \psi \wedge [\varphi!] \psi) \rightarrow \neg \varphi$

2 ...

by ...

by ...

- $\vdash_{PAL} [\varphi!] (\psi_1 \wedge \psi_2) \leftrightarrow ([\varphi!] \psi_1 \wedge [\varphi!] \psi_2)$

proof: $[\varphi!]$ is a normal modal operator

- $\vdash_{PAL} [\varphi!] K_i \psi \leftrightarrow (\neg \varphi \vee \neg [\varphi!] \psi)$

proof: use no forgetting and no learning axioms

Reducing *PAL* to *KT5*

- idea: apply equivalences, rewriting from left to right

$$[\varphi!] \psi \quad \leftrightarrow \quad (\neg \varphi \vee \psi) \quad \text{if } \varphi \text{ is atomic}$$

$$[\varphi!] \neg \psi \quad \leftrightarrow \quad (\neg \varphi \vee \neg [\varphi!] \psi)$$

$$[\varphi!] (\psi_1 \wedge \psi_2) \quad \leftrightarrow \quad ([\varphi!] \psi_1 \wedge [\varphi!] \psi_2)$$

$$[\varphi!] K_i \psi \quad \leftrightarrow \quad (\neg \varphi \vee \neg [\varphi!] \psi)$$

- effect:

- ▶ 'push down' announcement operators
- ▶ eliminate when a Boolean formula is attained

- exercises: reduce

- ▶ $[\rho!] K_1 p$
- ▶ $[\rho!] K_1 K_2 p$
- ▶ $[(p \wedge \neg K_1 p)!] K_1 p$

Reducing *PAL* to *KT5*, ctd.

- $red(\varphi)$ = result of reduction of φ

Reduction Theorem.

for every *PAL*-formula φ

(instance, not schema!)

$red(\varphi)$ is a *KT5*-formula, and $\vdash_{PAL} \varphi \leftrightarrow red(\varphi)$

Sketch of proof.

- equivalences are theorems
- substitution of proved equivalents (REq) preserves *PAL*-theoremhood
- define a decreasing counter (sum number of announcements governing subformulas)
 \Rightarrow rewriting terminates

Completeness theorem for *PAL*

Completeness Theorem.

$$\vdash_{PAL} \varphi \text{ iff } \models_{PAL} \varphi$$

Proof: ...

- *PAL* is decidable
- reduction to *KT5* leads to suboptimal decision procedure
- N.B.: rule of uniform substitution not *PAL*-valid:
 - ▶ $\vdash_{PAL} [p!]K_1 p$ (v.s.; p formula!)
 - ▶ $\not\vdash_{PAL} [\varphi!]K_i \varphi$ (v.s.; φ schema!)
- N.B.: the other axiom schemas of *K* are *PAL*-valid, too
 - ▶ reduction axioms suffice to prove all valid formula instances
but don't allow to prove all valid formula schemas

Muddy children reloaded

- *positive formula* π :

$$\pi ::= \beta \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid K_i \varphi$$

where β ranges over Boolean formulas

- prove that $\vdash_{PAL} \pi \rightarrow [\varphi!] \pi$ if π is a positive formula

- ▶ induction step for $\pi = K_i \pi_1$:

1 $\pi_1 \rightarrow [\varphi!] \pi_1$

2 $K_i \pi_1 \rightarrow K_i [\varphi!] \pi_1$

3 $K_i [\varphi!] \pi_1 \rightarrow [\varphi!] K_i \pi_1$

4 $K_i \pi_1 \rightarrow [\varphi!] K_i \pi_1$

by induction hyp.
by rule RM(K_i)
no forgetting
from 2. and 3. by CPL

- prove that $\vdash_{PAL} [\pi!] \pi$ if π is a positive formula

- ▶ $\vdash_{PAL} \pi \rightarrow [\pi!] \pi$ because ...

- ▶ $\vdash_{PAL} \neg \pi \rightarrow [\pi!] \pi$ because ...

- show:

- ▶ $\vdash_{PAL} [(m_1 \vee m_2)!] K_1 K_2 (m_1 \vee m_2)$

- ▶ $\vdash_{PAL} [\neg K_2 m_2!] K_1 \neg K_2 m_2$

- ▶ $\vdash_{KT5} (K_1 K_2 (m_2 \vee m_1) \wedge K_1 \neg K_2 m_2 \rightarrow K_1 \neg K_2 \neg m_1$

- ▶ $\vdash_{KT5} (K_1 \neg K_2 \neg m_1 \wedge K_1 (K_2 \neg m_1 \vee K_2 m_1)) \rightarrow K_1 K_2 m_1$

- conclude that

$$\vdash_{PAL} K_1 (K_2 \neg m_1 \vee K_2 m_1) \rightarrow [(m_1 \vee m_2)!] [\neg K_2 m_2!] K_1 m_1$$

Excursion: the paradox of knowability [Fitch]

- add a new modal operator quantifying over announcements:
 - ▶ $M, w \Vdash \Diamond\varphi$ iff there is ψ such that $M, w \Vdash \langle\psi\rangle\varphi$
 - ★ N.B.: ψ should have no occurrence of \Diamond (why?)
- allows to reason about plan existence (epistemic actions only)
 - ▶ $\models^? Init \rightarrow \Diamond Goal$
- example: $\models \Diamond(K_i p \vee K_i \neg p)$
- verificationist thesis:
 - ▶ $\varphi \rightarrow \Diamond K_i \varphi$ should be valid for every φ
- paradox of knowability:
 - ▶ $\not\models (p \wedge \neg K_i p) \rightarrow \Diamond K_i (p \wedge \neg K_i p)$

Moscow Mathematics Olympiad in 2000:

From a pack of seven known cards Ann and Bill each draw three cards and Cath gets the remaining card.

How can Ann and Bill publicly and truthfully inform each other about their cards, without Cath learning from any of their cards who holds it?

Moscow Mathematics Olympiad in 2000:

From a pack of seven known cards Ann and Bill each draw three cards and Cath gets the remaining card.

How can Ann and Bill publicly and truthfully inform each other about their cards, without Cath learning from any of their cards who holds it?

- does not work for any number and any distribution of cards
 - ▶ for which numbers there is a solution? (open problem)

Moscow Mathematics Olympiad in 2000:

From a pack of seven known cards Ann and Bill each draw three cards and Cath gets the remaining card.

How can Ann and Bill publicly and truthfully inform each other about their cards, without Cath learning from any of their cards who holds it?

- does not work for any number and any distribution of cards
 - ▶ for which numbers there is a solution? (open problem)
- perspective: unconditionally secure cryptographic protocols (perfect reasoners, public communication)
 - ▶ RSA: presupposes non-omniscience (decomposition into prime factors not feasible)

- cards are 0, . . . , 6, and Ann holds 012 and Bill holds 345
- some bad solutions:
 - ▶ Ann says: “Cath holds 6”
 - ★ Ann can only announce what she knows!
 - ▶ Ann says: “I don’t hold 6”
 - ★ Ann should know that Cath doesn’t learn anything!
 - ▶ Ann says: “I our Bill hold 012” (and Bill: “I our Ann hold 345”)
 - ★ Cath learns that Ann has 012!
 - ▶ Ann says: “either I hold 012, or I hold none of 0, 1, 2”
 - ★ Cath doesn’t learn any card,
 - ★ Ann knows that,
 - ★ but Cath does not know *that*

(that Cath remains ignorant it should be *common knowledge*)

- solutions:
 - ▶ Ann says: “My cards are among 012, 034, 056, 135 and 246”, and then Bill says: “Cath has 6”
 - ▶ . . .
- can be modelled in *PAL*

Dynamic epistemic logic *DEL*

- *PAL*: announcements are perceived by every agent:
 - ▶ $[p!](K_1 p \wedge K_2 p \wedge K_3 p \wedge \dots)$
- idea: epistemic logic models the agents' uncertainty about current state by means of possible states
⇒ model uncertainty about *current event* by possible events

static uncertainty

possible worlds

indistinguishability of worlds

dynamic uncertainty

possible events

indistinguishability of events

- example: suppose $p \wedge \neg K_1 p \wedge \neg K_1 \neg p \wedge \neg K_2 p \wedge \neg K_2 \neg p$
 - ▶ agent 2 learns that p
 - ▶ various possible perceptions of 1:
 - ★ 1 also learns that p , and 2 knows that, etc. ⇒ *PAL*
 - ★ 1 sees that 2 learns whether p , but does learn it himself (and 2 knows that, etc.)
 - ★ 1 does not see this (and 2 knows that, etc.)
 - ★ 1 *suspects this*
 - ★ ...

- static epistemic logic: static model $M^s = \langle W^s, R^s, V^s \rangle$
- dynamic epistemic logic: dynamic model $M^d = \langle W^d, R^d, V^d \rangle$,
where
 - ▶ W^d is a nonempty set of events
 - ▶ $R^d : Agts \longrightarrow W^d \times W^d$
 - ★ every R_i^d is an equivalence relation
 - ★ $eR_i e' = "i \text{ perceives occurrence of } e \text{ as occurrence of } e'"$
 - ▶ $V^d : W^d \longrightarrow Fmls$
 - ★ precondition of event w^d
- exercise: find dynamic models for the above examples

DEL: product construction

- given:

- ▶ a static model $M^s = \langle W^s, R^s, V^s \rangle$

- ▶ a dynamic model $M^d = \langle W^d, R^d, V^d \rangle$

what is the resulting static model?

DEL: product construction

- given:

- ▶ a static model $M^s = \langle W^s, R^s, V^s \rangle$
- ▶ a dynamic model $M^d = \langle W^d, R^d, V^d \rangle$

what is the resulting static model?

- $M = M^s \otimes M^d = \langle W, R, V \rangle$ where

- ▶ $W = \{ \langle w^s, w^d \rangle : w^s \in W^s, w^d \in W^d, \text{ and } M, w^s \Vdash V^d(w^d) \}$
- ▶ $R_i = \{ \langle \langle w^s, w^d \rangle, \langle v^s, v^d \rangle \rangle : w^s R_i^s v^s \text{ and } w^d R_i^d v^d \}$
- ▶ $V(\langle w^s, w^d \rangle) = V^s(w^s)$

restricted product

DEL: product construction

- given:

- ▶ a static model $M^s = \langle W^s, R^s, V^s \rangle$
- ▶ a dynamic model $M^d = \langle W^d, R^d, V^d \rangle$

what is the resulting static model?

- $M = M^s \otimes M^d = \langle W, R, V \rangle$ where

- ▶ $W = \{ \langle w^s, w^d \rangle : w^s \in W^s, w^d \in W^d, \text{ and } M, w^s \Vdash V^d(w^d) \}$
- ▶ $R_i = \{ \langle \langle w^s, w^d \rangle, \langle v^s, v^d \rangle \rangle : w^s R_i^s v^s \text{ and } w^d R_i^d v^d \}$
- ▶ $V(\langle w^s, w^d \rangle) = V^s(w^s)$

restricted product

- exercise: build outcome models for the above examples

- reduction axioms
- completeness (via reduction axioms)
- applications
 - ▶ Cluedo
 - ▶ cryptographic protocols
 - ▶ ...