Modal logics: applications and proof methods

Part I: Introduction to modal and multimodal logics

Andreas Herzig
IRIT-CNRS, Toulouse, France

www.irit.fr/~Andreas.Herzig
herzig@irit.fr

Nov. 2008
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
Chapter 1.
Motivation and introduction
write in the language of propositional logic:
- $\varphi_1 = \text{“Rare things are expensive”}$
- $\varphi_2 = \text{“Cheap things are rare”}$
- $\varphi_3 = \text{“Cheap things are expensive”}$

N.B.: stay propositional, i.e. avoid quantifiers and consider some arbitrary but fixed thing ‘$t$’; use atomic formulas $\text{Rare}_t$ and $\text{Expensive}_t$

prove that $\left(\varphi_1 \land \varphi_2\right) \rightarrow \varphi_3$ is valid in propositional logic

deduce formally that $\left(\varphi_1 \land \varphi_2\right) \rightarrow \varphi_3$ is a theorem of propositional logic

which piece of background knowledge (linguistic knowledge, alias ‘analytic proposition’) is not expressed in $\varphi_1, \varphi_2, \varphi_3$?
Warming up: predicate logic

- write in the language of first-order predicate logic:
  - $\varphi_1 = \text{"All humans are mortal"}$
  - $\varphi_2 = \text{"Socrates is a human"}$
  - $\varphi_3 = \text{"Socrates is mortal"}$

- deduce formally that $(\varphi_1 \land \varphi_2) \rightarrow \varphi_3$ is a theorem of predicate logic

- are there other possibilities to logically formulate $\varphi_1, \varphi_2, \varphi_3$?
  - which are better? (and what does ‘better’ mean here?)

- what are the main differences between propositional and predicate logic?
Warming up: arithmetics

- write in the language of predicate logic:
  - “0 is a natural number”
  - “if \( x \) is a natural number then \( \text{Succ}(x) \) is a natural number”
  - the induction axiom

  N.B.: use the unary predicate \( \text{Nat}(x) \) to express that \( x \) is a natural number

- what is the difference between first-order and second-order predicate logic?

- write the axioms for even and odd numbers
  N.B.: only use the function \( \text{Succ} \) and the predicate \( \text{Nat} \)
let $inc_x$ be an instruction (of some programming language) incrementing the value of program variable $x$

express in propositional or predicate logic:
- $\varphi =$ “if $x$ is even then after the execution of $inc_x$, $x$ is odd”

write a BNF for programs $\pi$, allowing for:
- atomic programs
- program composition (“;”)
- program iteration (“*”)
- testing the truth of a proposition (“?”)

work out the difference between programs and propositions

can you think of a way of writing this without referring to states?
constructive vs. non-constructive proofs

- express in predicate logic:
  - $\varphi = \text{“there are irrational numbers } x \text{ and } y \text{ such that } x^y \text{ is rational”}

- N.B.: use the language of predicate logic
  - unary predicate $Rat$
  - binary function $Power$
    - ... but for readability, write $x^y$ instead of $Power(x, y)$

- prove that $(Rat(2) \land \neg Rat(\sqrt{2}) \land Rat((\sqrt{2}^{\sqrt{2}})^{\sqrt{2}})) \rightarrow \varphi$ is a theorem of predicate logic
constructive vs. non-constructive proofs, ctd.

\[ \vdash_{\text{FOL}} \left( \text{Rat}(2) \land \neg \text{Rat}(\sqrt{2}) \land \text{Rat}((\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}) \right) \rightarrow \exists x \exists y (\neg \text{Rat}(x) \land \neg \text{Rat}(y) \land \text{Rat}(x^y)) \]

- reason by cases: prove that both \( \text{Rat}(\sqrt{2}^{\sqrt{2}}) \rightarrow \varphi \) and \( \neg \text{Rat}(\sqrt{2}^{\sqrt{2}}) \rightarrow \varphi \) are theorems
- non-constructive proof: doesn’t prove that \( \sqrt{2}^{\sqrt{2}} \) is irrational!
  - only proved in the 50ies
  - first constructive proof of \( \varphi \)
- constructive (‘intuitionistic’) mathematics
  - rejects axiom \( \varphi \lor \neg \varphi \) (‘tertium non datur’)
  - rejects axiom \( (\neg \varphi \rightarrow \bot) \rightarrow \varphi \) (‘reductio ad absurdum’)

A. Herzig ()

Modal logics

Nov. 2008 9 / 35
Reasoning about knowledge

‘knowing that there is a number’ vs. ‘knowing the number’

- write in the language of predicate logic:
  - “Hilbert knows that there are irrational $x$ and $y$ such that $x^y$ is rational”
  - “there are irrational $x$ and $y$ such that Hilbert knows that $x^y$ is rational”
  - “there are irrational $x$ and $y$ such that $x^y$ is rational, but Hilbert does not know that”

- hint: for readability, you may abbreviate $\neg \text{Rat}(x) \land \neg \text{Rat}(y) \land \text{Rat}(x^y)$ by $P(x, y)$

- N.B.: as Hilbert knew the axioms PA of Peano Arithmetic, he should have been able to prove that $\exists x \exists y P(x, y)$
  - … if he was a perfect, ‘omniscient’ reasoner
  - to be discussed later
Reasoning about knowledge: muddy children

a famous puzzle:

1. two children come back from the garden, both with mud on their forehead; their father looks at them and says:
   “at least one of you has mud on his forehead”
   then he asks:
   “those who know whether they are dirty, step forward!”
2. nobody steps forward
3. the father asks again:
   “those who know whether they are dirty, step forward!”
4. both simultaneously answer: “I know!”

N.B.: can be generalized to an arbitrary number $n \geq 2$ of children
use (second-order) predicate $\text{Knows}(i, \varphi)$, where $i \in \{1, 2\}$

- $\text{Knows}(i, \varphi) = \text{“agent } i \text{ knows that } \varphi\text{”}$

some of child 2’s knowledge at the different stages:

(S0) background knowledge:
$\text{Knows}(2, \text{Knows}(1, m_2) \lor \text{Knows}(1, \neg m_2))$

equivalently:
$\text{Knows}(2, \neg \text{Knows}(1, \neg m_2) \rightarrow \text{Knows}(1, m_2))$

(S1) learns that at least one of them has mud on his forehead:
$\text{Knows}(2, \text{Knows}(1, (m_1 \lor m_2)))$

(S2) child 2 does not respond:
$\text{Knows}(2, \neg \text{Knows}(1, m_1))$

(S3) should follow from (S0)-(S2):
$\text{Knows}(2, m_2)$

proof?
Reasoning about knowledge: muddy children

deduction of (S3) from (S0), (S1), (S2):

1. \( \text{Knows}(2, \text{Knows}(1, (m_1 \lor m_2))) \)  
   hyp. (S1)
2. \( \text{Knows}(2, \text{Knows}(1, \neg m_2) \rightarrow \text{Knows}(1, m_1)) \)  
   conseq. of 1.
3. \( \text{Knows}(2, \neg \text{Knows}(1, m_1) \rightarrow \neg \text{Knows}(1, \neg m_2)) \)  
   equiv. to 2.
4. \( \text{Knows}(2, \neg \text{Knows}(1, m_1)) \)  
   hyp. (S2)
5. \( \text{Knows}(2, \neg \text{Knows}(1, \neg m_2)) \)  
   from 3. and 4.
6. \( \text{Knows}(2, \neg \text{Knows}(1, \neg m_2) \rightarrow \text{Knows}(1, m_2)) \)  
   equiv. to hyp. (S0)
7. \( \text{Knows}(2, \text{Knows}(1, m_2)) \)  
   from 5. and 6.
8. \( \text{Knows}(2, m_2) \)  
   from 7., bec. \( \text{Knows}(1, m_2) \rightarrow m_2 \)  
   (‘knowledge implies truth’)

informal deduction \( \Rightarrow \) formal rules? \( \Rightarrow \) deduction in a formal logic?
A second-order theory of the \textit{Knows} predicate

- **desirable principles:**
  - $\forall i \forall p \left( \text{Knows}(i, p) \rightarrow p \right)$
    - $\star$ used in step 8.
  - $\forall i \forall p \forall q \left( \text{Knows}(i, p) \land \text{Knows}(i, p \rightarrow q) \rightarrow \text{Knows}(i, q) \right)$
    - $\star$ used in step 2.
  - $\ldots$

- **make up theory of knowledge $\mathcal{T}_{\text{Knows}}$**
  - second-order formulas ("$\forall p$" quantifies over propositions)

- **reasoning about knowledge:**
  - $\mathcal{T}_{\text{Knows}} \vdash ((S0) \land (S1) \land (S2)) \rightarrow (S3)$
  - consequence problem in second-order logic
    - $\star$ undecidable $\ldots$
**Knows: from second-order to first-order logic**

**idea [Hintikka 62]:**

\[ \text{Knows}(i, \varphi) = \text{“} \varphi \text{ true in all worlds that are possible for } i \text{”} \]

- set of possible worlds \( W \)
- ternary ‘accessibility’ relation \( R(i, w_1, w_2) \)
  - \( i = \text{agent} \)
  - \( w_1 = \text{actual world} \)
  - \( w_2 = \text{world that } i \text{ cannot distinguish from } w_1 \)

**in first-order logic:**

\[ \text{Knows}(i, \varphi, w) = \text{“} \text{at } w, i \text{ knows that } \varphi \text{”} \]

\[ \begin{align*}
\text{def} & \quad \forall w' \ (R(i, w, w') \rightarrow \varphi[w']) \\
\equiv & \quad \forall w' \ (R(i, w, w') \rightarrow \varphi[w'])
\end{align*} \]
muddy children:

- \( \text{Knows}(1, m_2, w) = \forall w' \ (R(1, w, w') \rightarrow m_2(w')) \)
- \( \neg \text{Knows}(1, m_1, w) = \exists w' \ (R(1, w, w') \land \neg m_1(w')) \)

draw the set of possible worlds and the accessibility relation

- in the initial situation
- after the father has announced \( m_1 \lor m_2 \)
- after the first round (when none of the children stepped forward)
desirable principles for knowledge $\Rightarrow$ properties of $R$

- $\forall i \forall p \ (\text{Knows}(i, p) \rightarrow p)$ corresponds to: $\forall i \forall w \ R(i, w, w)$
- ...

make up first-order theory $\mathcal{T}_{\text{Knows}}$

reasoning about knowledge:

- $\mathcal{T}_{\text{Knows}} \vdash \forall w ((S_0 \wedge S_1 \wedge S_2) \rightarrow (S_3))[w]$
- consequence problem in first-order logic
  - semi-decidable ...
**Knows: from first-order to modal logic**

**idea [Hintikka 62]:**

Don’t use first-order language, but add **modal operators of knowledge** to the language of propositional logic

- \( K_i \) = modal operator (modifies the sense of propositions)
- **epistemic** language:

\[
\varphi ::= p \mid \neg \varphi \mid (\varphi \land \varphi) \mid K_i \varphi
\]

where \( p \) ranges over set of propositional atoms \( Atms \) and \( i \) over the set of agents \( Agts \)

- Reading: \( K_i \varphi = “i \text{ knows that } \varphi” \)
- N.B.: propositional language
  - No quantifiers \( \forall, \exists \)
Epistemic language: examples

- **knowing-whether:**
  - $K_1 m_2 \lor K_1 \neg m_2$  
  child 1 knows whether $m_2$

- **ignorance:**
  - $\neg K_2 m_2 \land \neg K_2 \neg m_2$  
  child 2 does not know whether $m_2$

- **abbreviation:**
  - $\hat{K}_i \varphi \overset{\text{def}}{=} \neg K_i \neg \varphi = \text{“}\varphi\text{ is possible for } i\text{”}$

- **nesting of modal operators (‘higher-order knowledge’):**
  - $K_2 (\neg K_1 m_1 \land (K_1 m_2 \lor K_1 \neg m_2))$
The propositional logic of knowledge $EL$

- extend propositional logic by axiom schemas and inference rules for $K_i$
  - $\vdash_{EL} K_i \varphi \rightarrow \varphi$
  - if $\vdash_{EL} \varphi$ then $\vdash_{EL} K_i \varphi$
  - ...

- logic of knowledge = epistemic logic $EL$

- reasoning about knowledge:
  - $\vdash_{EL} K_2 K_1 m_2 \rightarrow K_2 m_2$
  - $\vdash_{EL} ((S0) \land (S1) \land (S2)) \rightarrow (S3)$
  - ...
  - reasoning problem: given $\varphi$, do we have $\vdash_{EL} \varphi$?
    - decidable!
    - PSPACE complete (propositional logic: NP complete)
    - more details later ...
Reasoning in epistemic logic *EL*

- **Semantics:** models? truth conditions?
  - resort to first-order semantics in terms of possible worlds
  - \( M = \langle W, R, V \rangle \) where
    - \( W \) some set (‘possible worlds’)
    - \( R : \text{Agts} \times W \times W \)
    - \( V \) valuation
  - truth conditions:
    - \( M, w \models K_i \varphi \iff M, w' \models K_i \varphi \) for all \( w' \) such that \( R(i, w, w') \)
  - N.B.: language of *EL* is less expressive than that of *FOL*
A generalization: modalities

- $\text{After}_a$ and $K_i$ are **modalities**
  - modify the sense of propositions
- modalities are not truth functional
  - truth value of $K_i \varphi$ is not function of truth value of $\varphi$
  - remember: $\neg, \land, \lor, \rightarrow$ are truth functional
  - $\Rightarrow$ models should consist of more than just a valuation function
A generalization: modalities, ctd.

- other modalities:
  - always $m_1 \rightarrow K_2 m_1$ (temporal)
  - sometimes $m_1 \land \neg K_2 m_1$ (temporal)
  - $2$ believes that $m_2$ (doxastic)
  - it is probable for $2$ that $m_2$ (doxastic)
  - $i$ wants $\neg m_2$ (intentional)
  - it is obligatory that $\neg m_2$ (deontic)
  - it is permitted that $m_2$ (deontic)
  - it is forbidden that $m_2$ (deontic)
  - necessarily $m_1 \land m_2$ (alethic)
  - possibly $\neg m_1 \lor \neg m_2$ (alethic)
  - ...

- modalities can be combined:
  - $2$ believes that $1$ knows whether $m_2$
  - it is always the case that $\neg m_2$ after cleaning
  - $2$ knows that after cleaning, $\neg m_2$
  - ...

Modalities are useful

- modalities do not occur in mathematical reasoning
  - exception: the concept of provability in arithmetic [Gödel 32]
- but are central in:
  - program verification
  - intelligent agent specification
  - multi-agent systems design
  - commonsense reasoning
  - semantics of natural language
  - cognitive economy
  - ...

⇒ uniform analysis?
Dual modalities

- dual modalities: universal / existential
  - always / sometimes
  - obligatory / permitted
  - necessarily / possibly
  - ...

- generic modal operators:
  - $\Box_i = \text{‘necessarily’ (universal)}$
  - $\Diamond_i = \text{‘possibly’ (existential)}$
  - $i = \text{parameter}$
    - agent / program / normative system / ...
    - should allow to distinguish the different operators under concern
    - also used: $\Box_i = [i]$ and $\Diamond_i = \langle i \rangle$

- duality:
  - $\Box_i \varphi \leftrightarrow \neg \Diamond_i \neg \varphi$ and $\Diamond_i \varphi \leftrightarrow \neg \Box_i \neg \varphi$
  - $\Box_i$ and $\Diamond_i$ interdefinable
  - here: $\Box_i$ primitive, and $\Diamond_i \varphi$ abbreviates $\neg \Box_i \neg \varphi$
Modalities and their logics

- uniform semantics: ‘possible worlds models’ [Kripke 59]
  - set of possible worlds
  - accessibility relations

  ⇒ normal modal logics

- varying properties of the accessibility relations
  - reflexive, transitive, symmetric, serial, dense, . . . , confluent, inclusion, . . .

- properties of $R_i$ correspond to properties of $\Box_i$
  - if $\Box_i$ is epistemic then $R(\Box_i)$ has to be reflexive
  - if $\Box_i$ is doxastic then $R(\Box_i)$ has to be serial (but not necessarily reflexive)

- relations between modalities
  - doxastic relation contained in epistemic relation:
    - if $\Box_{K_i}$ is epistemic and $\Box_{B_i}$ is doxastic then $R(\Box_{B_i}) \subseteq R(\Box_{K_i})$
    - guarantees that knowledge implies belief

Modalities and their logics, ctd.

- useful? fruitful?
  - new questions, new problems? links with other formalisms?
- range of applicability? limitations?
  - e.g. omniscience problem in epistemic logics
  - computational costs
- mathematical analysis:
  - soundness?
  - completeness?
  - decidability?
  - complexity of satisfiability?
Recap of basic logic notions: language

- **primitive symbols:**
  - logical symbols: \( \neg, \wedge, \perp, \rightarrow, \ldots, K, B, \text{After}, \ldots, \Box, \Diamond, \ldots, \)
  - sets of non-logical symbols:
    - set of propositional atoms \( \text{Atms} = \{ p, q, \ldots \} \)
    - set of agents \( \text{Agts} = \{ i, j, \ldots \} \)
    - set of atomic actions \( \text{Acts} = \{ a, b, \ldots \} \)
    - \( \ldots \)
  - parentheses

- **language** = set of formulas, defined from primitive symbols by means of a grammar
  - Backus-Naur-form (BNF):
    \[
    \varphi ::= p \mid \perp \mid \neg \varphi \mid (\varphi \wedge \varphi) \mid (\varphi \rightarrow \varphi) \mid \ldots \mid K_i \varphi \mid B_i \varphi \mid \text{After}_a \varphi \mid \ldots
    \]
Recap of basic logic notions: subformulas

- inductive definition of the set of subformulas $sf(\varphi)$ of $\varphi$:
  
  \[
  \begin{align*}
  sf(p) & = \{p\} \\
  sf(\neg \varphi) & = sf(\varphi) \cup \{\neg \varphi\} \\
  sf(\varphi \land \psi) & = sf(\varphi) \cup sf(\psi) \cup \{\varphi \land \psi\} \\
  sf(\Box_i \varphi) & = sf(\varphi) \cup \{\Box_i \varphi\}
  \end{align*}
  \]

- suppose $\Box_i \psi \in sf(\varphi)$
  
  - in $\varphi$, $\chi$ is in the scope of $\Box_i$ iff $\chi \in sf(\psi)$
  - in $\varphi$, $\Box_j$ is in the scope of $\Box_i$ iff $\Box_j \chi \in sf(\psi)$ for some $\chi$
Recap of basic logic notions

- **logic** $\Lambda = \text{language } \mathcal{L}_{\Lambda} + \text{particular subset of } \mathcal{L}_{\Lambda}$ (called theorems or validities)

- **particular subset** of $\mathcal{L}_{\Lambda}$ can be characterized in two ways:
  - semantically: using models $\Rightarrow$ validities
  - syntactically: using axioms and inference rules $\Rightarrow$ theorems
Recap of basic logic notions: axiomatics

- requires:
  1. **axiom schemas** = basic theorems of the logic
     - in an axiom schema, we can perform *uniform substitutions*:
       \[ K_i \varphi \rightarrow \varphi \text{ instantiates to: } K_1 (m_2 \lor m_1) \rightarrow (m_2 \lor m_1) \]
     - N.B.: the \( \varphi \) are *meta-variables* over the language
  2. **inference rules** = generate new theorems from existing theorems
     - notation: \( \{ \varphi_1, \ldots, \varphi_m \} / \varphi \), or: \( \frac{\varphi_1, \ldots, \varphi_m}{\varphi} \)

- a **proof** of \( \varphi \) in \( \Lambda \) is a sequence of formulas \( \langle \varphi_1, \ldots, \varphi_n \rangle \) such that \( \varphi_n = \varphi \), and for every \( i \leq n \):
  - \( \varphi_i \) is an (instance of) some axiom schema for \( \Lambda \), or
  - there are formulas \( \varphi_{i_1}, \ldots, \varphi_{i_m} \), such that \( i_j < i \), and \( \frac{\varphi_{i_1}, \ldots, \varphi_{i_m}}{\varphi_i} \) is (an instance of) some inference rule for \( \Lambda \)

- \( \varphi \) is a **theorem** of \( \Lambda \) iff \( \varphi \) is provable in \( \Lambda \)
  - notation: \( \vdash_\Lambda \varphi \)

- \( \varphi \) is **consistent** in \( \Lambda \) iff \( \nvdash_\Lambda \neg \varphi \)

- **deductions** \( \Gamma \vdash_\Lambda \varphi \) iff . . . (several options in modal logic, v.i.)
Recap of basic logic notions: semantics

requires:

1. a class of models $M$ for $\Lambda$
2. truth conditions: when is $\phi$ true in $M$?
   * notation in general: $M \vDash \phi$
   * in modal logic: $M, w \vDash \phi$

$\phi$ is valid in $\Lambda$ iff $M, w \vDash \phi$, for every model $M$ for $\Lambda$ and world $w$ in $M$
   * notation: $\models_{\Lambda} \phi$

$\phi$ is satisfiable in $\Lambda$ iff $\not\models_{\Lambda} \neg \phi$

logical consequence $\Gamma \models_{\Lambda} \phi$ iff . . . (several options in modal logic, v.i.)
Recap of basic logic notions: soundness and completeness

Syntactic and semantic characterizations should coincide!

**soundness**: for every formula \( \varphi \), if \( \vdash \Lambda \varphi \) then \( \models \Lambda \varphi \)
- proof by induction on the length of the proof of \( \varphi \)
  - base: every instance of every axiom schema is valid
  - step: every inference rule preserves validity

**completeness**: for every formula \( \varphi \), if \( \models \Lambda \varphi \) then \( \vdash \Lambda \varphi \)
- actually proved: ‘if \( \varphi \) is consistent in \( \Lambda \) then \( \varphi \) is satisfiable in \( \Lambda \)’
  - implies: ‘if \( \neg \varphi \) is consistent in \( \Lambda \) then \( \neg \varphi \) is satisfiable in \( \Lambda \)’
  - which is equivalent to: ‘if \( \not \models \Lambda \neg \neg \varphi \) then \( \not \models \Lambda \neg \varphi \)’
  - which is equivalent to: ‘if \( \not \models \Lambda \varphi \) then \( \not \models \Lambda \neg \varphi \)’
- non-constructive proofs: canonical models [Henkin]
- constructive proofs: via tableau method

**strong completeness**: if \( \Gamma \models \Lambda \varphi \) then \( \Gamma \vdash \Lambda \varphi \)
- implies weak completeness, but not other way round
Yet another motivation to study modal logic

- idea: explore interval between classical propositional logic (CPL) and first-order logic (FOL)
  - “stay decidable, but express more”
  - more mathematical motivation

- decidable FOL fragments:
  - no quantifiers, no variables: CPL
  - only unary predicates
    - no dependencies between quantifiers
  - only universally quantified variables:
    - no function symbols
    - \{∀x_1 \ldots ∀x_n \varphi : \text{neither quantifiers occur in } \varphi, \text{ and every free variable in } \varphi \text{ is among } x_1, \ldots, x_n \}\n  - …
Yet another motivation to study modal logic, ctd.

- an large decidable fragment: guarded quantification
  - basic idea:
    for every formula $\varphi$ and every subformula $\forall y \psi$ of $\varphi$, $\psi = (R(i, x, y) \rightarrow \chi[y])$ and $y$ is the only free variable in $\chi$
    - cf. first-order formulation of modalities
  - description logics
    - family of knowledge representation languages, aka ‘terminological logics’: $\mathcal{AL}$, $\mathcal{ALC}$, ...
    - at the base of the semantic web: $\textit{OWL}$, $\textit{OWL}$-DL, $\textit{OWL}$-lite, ...
    - basic description logic $\mathcal{ALC} = \text{multimodal } K$ (v.i.)