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

Warming up: propositional logic

- write in the language of propositional logic:
  - φ_1 = “Rare things are expensive”
  - φ_2 = “Cheap things are rare”
  - φ_3 = “Cheap things are expensive”

  N.B.: stay propositional, i.e. avoid quantifiers and consider some arbitrary but fixed thing ‘t’; use atomic formulas $Rare_t$ and $Expensive_t$
- prove that $(\varphi_1 \land \varphi_2) \to \varphi_3$ is valid in propositional logic
- deduce formally that $(\varphi_1 \land \varphi_2) \to \varphi_3$ 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:
  - φ_1 = "All humans are mortal"
  - φ_2 = "Socrates is a human"
  - φ_3 = "Socrates is mortal"
- deduce formally that $(\varphi_1 \land \varphi_2) \to \varphi_3$ is a theorem of predicate logic
- are there other possibilities to logically formulate φ_1, φ_2, φ_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 $Succ(x)$ is a natural number"
  - the induction axiom

  N.B.: use the unary predicate $Nat(x)$ 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 *Succ* and the predicate *Nat*

Reasoning about events, actions and programs

let inc_x be an instruction (of some programming language) incrementing the value of program variable x

- express in propositional or predicate logic:
  - φ = "if x is even then after the execution of inc_x, x is odd"
- write a BNF for programs π, 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:
  - φ = "there are irrational numbers x and y such that x^y 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 $(Rat(2) \wedge \neg Rat(\sqrt{2}) \wedge Rat((\sqrt{2}^{\sqrt{2}})^{\sqrt{2}})) \rightarrow \varphi$ is a theorem of predicate logic

constructive vs. non-constructive proofs, ctd.

$$\vdash_{FOL} (Rat(2) \land \neg Rat(\sqrt{2}) \land Rat((\sqrt{2}^{\sqrt{2}})^{\sqrt{2}})) \to \exists x \exists y (\neg Rat(x) \land \neg Rat(y) \land Rat(x^y))$$

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

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 $\neg Rat(x) \wedge \neg Rat(y) \wedge Rat(x^y)$ 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

Reasoning about knowledge: muddy children

- use (second-order) predicate $Knows(i, \varphi)$, where $i \in \{1, 2\}$
  - $Knows(i, \varphi)$ = "agent i knows that φ"
- some of child 2's knowledge at the different stages:

(S0) background knowledge:
$Knows(2,\ Knows(1, m_2) \vee Knows(1, \neg m_2))$
equivalently:
$Knows(2,\ \neg Knows(1, \neg m_2) \rightarrow Knows(1, m_2))$

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

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

(S3) should follow from (S0)-(S2):
$Knows(2, m_2)$

- proof?

Reasoning about knowledge: muddy children

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

1. $Knows(2, \; Knows(1, (m_1 \vee m_2)))$ — hyp. (S1)
2. $Knows(2, \; Knows(1, \neg m_2) \rightarrow Knows(1, m_1))$ — conseq. of 1.
3. $Knows(2, \; \neg Knows(1, m_1) \rightarrow \neg Knows(1, \neg m_2))$ — equiv. to 2.
4. $Knows(2, \; \neg Knows(1, m_1))$ — hyp. (S2)
5. $Knows(2, \; \neg Knows(1, \neg m_2))$ — from 3. and 4.
6. $Knows(2, \; \neg Knows(1, \neg m_2) \rightarrow Knows(1, m_2))$ — equiv. to hyp. (S0)
7. $Knows(2, \; Knows(1, m_2))$ — from 5. and 6.
8. $Knows(2, \; m_2)$ — from 7., bec. $Knows(1, m_2) \rightarrow m_2$ ('knowledge implies truth')

informal deduction $\Rightarrow$ formal rules? $\Rightarrow$ deduction in a formal logic?

A second-order theory of the *Knows* predicate

- desirable principles:
  - $\forall i \forall p \, (Knows(i, p) \rightarrow p)$
    - used in step 8.
  - $\forall i \forall p \forall q \, (Knows(i, p) \wedge Knows(i, p \rightarrow q) \rightarrow Knows(i, q))$
    - used in step 2.
  - ...
- make up theory of knowledge $\mathcal{T}_{Knows}$
  - second-order formulas ("$\forall p$" quantifies over propositions)
- reasoning about knowledge:
  - $T_{Knows} \vdash ((S0) \wedge (S1) \wedge (S2)) \rightarrow (S3)$
  - consequence problem in second-order logic
    - undecidable ...

Knows: from second-order to first-order logic

idea [Hintikka 62]:

$Knows(i, \varphi)$ = "φ true in all worlds that are possible for i"

- set of possible worlds W
- ternary 'accessibility' relation $R(i, w_1, w_2)$
  - i = agent
  - w_1 = actual world
  - w_2 = world that i cannot distinguish from w_1
- in first-order logic:

$$Knows(i, \varphi, w) \quad = \quad \text{"at } w, i \text{ knows that } \varphi\text{"}$$

$$\stackrel{\text{def}}{=} \quad \forall w' \, (R(i, w, w') \rightarrow \varphi[w'])$$

Knows: from second-order to first-order logic, ctd.

- muddy children:
  - $Knows(1, m_2, w) \;=\; \forall w' \; (R(1, w, w') \rightarrow m_2(w'))$
  - $\neg Knows(1, m_1, w) \;=\; \exists w' \; (R(1, w, w') \wedge \neg m_1(w'))$
- draw the set of possible worlds and the accessibility relation
  - in the initial situation
  - after the father has announced $m_1 \vee m_2$
  - after the first round (when none of the children stepped forward)

Knows: from second-order to first-order logic, ctd.

- desirable principles for knowledge $\Rightarrow$ properties of R
  - $\forall i \forall p\ (Knows(i,p) \to p)$ corresponds to: $\forall i \forall w\ R(i,w,w)$
  - ...
- make up first-order theory $\mathcal{T}_{Knows}$
- reasoning about knowledge:
  - $T_{Knows} \vdash \forall w\ (((S0) \land (S1) \land (S2)) \to (S3))[w]$
  - 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:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathrm{K}_i\, \varphi$$

where p ranges over set of propositional atoms *Atms* and i over the set of agents *Agts*

- reading: $\mathrm{K}_i\, \varphi$ = "i knows that φ"
- N.B.: propositional language
  - no quantifiers $\forall$, $\exists$

Epistemic language: examples

- knowing-whether:
  - $K_1\, m_2 \vee K_1\, \neg m_2$ child 1 knows whether m_2
- ignorance:
  - $\neg K_2\, m_2 \wedge \neg K_2\, \neg m_2$ child 2 does not know whether m_2
- abbreviation:
  - $\hat{K}_i\, \varphi \stackrel{\text{def}}{=} \neg K_i\, \neg \varphi$ = "φ is possible for i"
- nesting of modal operators ('higher-order knowledge'):
  - $K_2\, (\neg K_1\, m_1 \wedge (K_1\, m_2 \vee K_1\, \neg m_2))$

The propositional logic of knowledge *EL*

- extend propositional logic by axiom schemas and inference rules for K_i
  - $\vdash_{EL} \mathrm{K}_i \varphi \to \varphi$
  - if $\vdash_{EL} \varphi$ then $\vdash_{EL} \mathrm{K}_i \varphi$
  - ...
- **logic** of knowledge = epistemic logic *EL*
- reasoning about knowledge:
  - $\vdash_{EL} \mathrm{K}_2 \mathrm{K}_1 m_2 \to \mathrm{K}_2 m_2$
  - $\vdash_{EL} ((S0) \land (S1) \land (S2)) \to (S3)$
  - ...
  - reasoning problem: given φ, 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')
    - $R : Agts \times W \times W$
    - V valuation
  - truth conditions:
    - $M, w \Vdash \mathrm{K}_i\, \varphi$ iff $M, w' \Vdash \mathrm{K}_i\, \varphi$ for all w' such that $R(i, w, w')$
  - N.B.: language of *EL* is less expressive than that of *FOL*

A generalization: modalities

- After_a and K_i are **modalities**
  - modify the sense of propositions
- modalities are not truth functional
  - truth value of $\text{K}_i\,\varphi$ is not function of truth value of φ
  - remember: $\neg$, $\wedge$, $\vee$, $\rightarrow$ are truth functional

$\Rightarrow$ models should consist of more than just a valuation function

A generalization: modalities, ctd.

- other modalities:
  - always $m_1 \to K_2\, m_1$ (temporal)
  - sometimes $m_1 \wedge \neg K_2\, m_1$ (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 $m_1 \wedge m_2$ (alethic)
  - possibly $\neg m_1 \vee \neg m_2$ (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
  - ...

$\Rightarrow$ uniform analysis?

Dual modalities

- dual modalities: universal / existential
  - always / sometimes
  - obligatory / permitted
  - necessarily / possibly
  - ...
- generic modal operators:
  - $\Box_i$ = ‘necessarily’ (universal)
  - $\Diamond_i$ = ‘possibly’ (existential)
  - i = 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

$\Rightarrow$ **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: $\neg, \wedge, \bot, \rightarrow, \ldots,$ K , B , After , $\ldots, \Box, \Diamond, \ldots,$
  - sets of non-logical symbols:
    - ★ set of propositional atoms $Atms = \{p, q, \ldots\}$
    - ★ set of agents $Agts = \{i, j, \ldots\}$
    - ★ set of atomic actions $Acts = \{a, b, \ldots\}$
    - ★ ...
  - parentheses
- **language** = set of formulas, defined from primitive symbols by means of a grammar
  - Backus-Naur-form (BNF):

$$\varphi \quad ::= \quad p \mid \bot \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \rightarrow \varphi) \mid \ldots \mid \mathrm{K}_i\,\varphi \mid \mathrm{B}_i\,\varphi \mid \mathrm{After}_a\,\varphi \mid \ldots$$

Recap of basic logic notions: subformulas

- inductive definition of the set of subformulas $\mathit{sf}(\varphi)$ of φ:

$$
\begin{aligned}
\mathit{sf}(p) &= \{p\} \\
\mathit{sf}(\neg\varphi) &= \mathit{sf}(\varphi) \cup \{\neg\varphi\} \\
\mathit{sf}(\varphi \wedge \psi) &= \mathit{sf}(\varphi) \cup \mathit{sf}(\psi) \cup \{\varphi \wedge \psi\} \\
\mathit{sf}(\Box_i\varphi) &= \mathit{sf}(\varphi) \cup \{\Box_i\varphi\}
\end{aligned}
$$

- suppose $\Box_i\psi \in \mathit{sf}(\varphi)$
  - in φ, χ is in the *scope* of $\Box_i$ iff $\chi \in \mathit{sf}(\psi)$
  - in φ, $\Box_j$ is in the scope of $\Box_i$ iff $\Box_j\chi \in \mathit{sf}(\psi)$ for some χ
  -

Recap of basic logic notions

- **logic** Λ = language $\mathcal{L}_\Lambda$ + *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*: $K_i \varphi \to \varphi$ instantiates to: $K_1 (m_2 \vee m_1) \to (m_2 \vee m_1)$
     - N.B.: the φ 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 φ in Λ is a sequence of formulas $\langle \varphi_1, \ldots, \varphi_n \rangle$ such that $\varphi_n = \varphi$, and for every $i \leq n$:
  - φ_i is an (instance of) some axiom schema for Λ, 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 Λ
- φ is a **theorem** of Λ iff φ is provable in Λ
  - notation: $\vdash_\Lambda \varphi$
- φ is **consistent** in Λ iff $\not\vdash_\Lambda \neg\varphi$
- **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 Λ
  2. **truth conditions**: when is φ true in M?
     - notation in general: $M \Vdash \varphi$
     - in modal logic: $M, w \Vdash \varphi$ ‘φ is true in $\langle M, w \rangle$’
- φ is **valid** in Λ iff $M, w \Vdash \varphi$, for every model M for Λ and world w in M
  - notation: $\models_\Lambda \varphi$
- φ is **satisfiable** in Λ iff $\not\models_\Lambda \neg\varphi$
- **logical consequence** $\Gamma \models_\Lambda \varphi$ 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 φ, if $\vdash_\Lambda \varphi$ then $\models_\Lambda \varphi$
  - proof by induction on the length of the proof of φ
    - base: every instance of every axiom schema is valid
    - step: every inference rule preserves validity
- **completeness**: for every formula φ, if $\models_\Lambda \varphi$ then $\vdash_\Lambda \varphi$
  - actually proved: 'if φ is consistent in Λ then φ is satisfiable in Λ'
    - implies: 'if $\neg\varphi$ is consistent in Λ then $\neg\varphi$ is satisfiable in Λ'
    - which is equivalent to: 'if $\not\vdash_\Lambda \neg\neg\varphi$' then $\not\models_\Lambda \neg\neg\varphi$'
    - which is equivalent to: 'if $\not\vdash_\Lambda \varphi$ then $\not\models_\Lambda \varphi$'
  - 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 (monadic fragment)
    - no dependencies between quantifiers
  - only universally quantified variables:
    - no function symbols
    - $\{\forall x_1 \ldots \forall x_n \varphi \ : \ $ neither quantifiers occur in φ, and every free variable in φ is among $x_1, \ldots, x_n\}$
  - ...

Yet another motivation to study modal logic, ctd.

- an large decidable fragment: guarded quantification
  - basic idea:
    for every formula φ and every subformula $\forall y \psi$ of φ,
    $\psi = (R(i, x, y) \rightarrow \chi[y])$ and y is the only free variable in χ
    - 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: $\mathcal{OWL}$, $\mathcal{OWL}$-DL, $\mathcal{OWL}$-lite, ...
    - basic description logic $\mathcal{ALC}$ = multimodal K (v.i.)