Multiagent Systems

15. Logics for Multiagent System

B. Nebel, C. Becker-Asano, S. Wölfl

Albert-Ludwigs-Universität Freiburg

July 25, 2014

Where are we?

Last time . . .

- Argumentation: a richer form of negotiation
- Logic-based negotiation: attacks, defeats
- Strengths of arguments
- Abstract argumentation systems
- Argumentation dialogue systems

Today:

- Logics for Multiagent Systems

15.1 Introduction

Logics for multiagent systems

- Throughout computer science, logic is used to develop formal models of computation
- In multiagent systems, the predominant approach for doing this is based on modal logics
- These are used to model agents' belief states (but also other approaches, e.g. modelling commitments, obligations and permissions, etc.)
- We will first introduce the most common model of modal logic semantics, then use it to model beliefs and knowledge

Why modal logic?

We are looking for a logic that governs rational belief.
For example, if an agent believes A and if B follows logically from A, it is rational for the agent to believe B.

Do we need a **new** logic?

- Consider the following statement: "Michael believes that Ann likes the MAS course"
- Naive attempt: use first-order logic to express this as:

$$\mathrm{Bel}(\mathit{Michael}, \mathit{Likes}(\mathit{Ann}, \mathit{MAS}))$$

- ... but this is not a syntactically correct FOL formula (formulae cannot be used as terms)!
- We could introduce terms for statements like "Likes(Ann, MAS)", but that would not help very much: we would not be able to draw conclusions about what is believed by Michael

Why modal logic?

The semantic problem is even worse:

- If Ann has a daughter Mary, it holds $Ann = mother(Mary)$
- But would we conjecture that

$$\text{Bel}(Michael, Likes(mother(Mary), MAS))?$$

  After all, Michael might not know about this equality ...
- **Problem**: intentional notions are **referentially opaque**, they set up opaque contexts in which FOL substitution rules don't apply
- Classical logic based on **truth-functional** operators: the truth value of $p \wedge q$ is a function of the truth values of p and q
- Semantic value (denotation) of a formula depends only on denotations of sub-expressions
- But "Michael believes p" is not truth-functional, it depends on truth value of p **and** Michael's belief

15.2 Modal logics

Possible-worlds semantics

- Kripke's (1963) model of possible worlds: standard semantics in modal logics
- Example: a game of cards, agents cannot see each others set of cards
  - useful for agent to infer which cards are held by others
  - consider all possible distributions of cards among all players
  - own cards (and cards on the table) eliminate certain distributions
  - each remaining possible distribution of the cards is a **possible world**
- We can describe the agents belief by the set of worlds he thinks possible: **epistemic alternatives**

Normal modal logic

- Before moving to epistemic logic we describe the framework of **normal modal logic** as its foundation
- Based on distinction between necessary and contingent truths
- Necessary truths are true in all possible worlds, possible truths are true in some possible worlds
- Use $\Box$ (**box**) and $\Diamond$ (**diamond**) operators read as: "necessarily" and "possibly", respectively
- We introduce a simple **propositional modal** logic (classical propositional logic extended with the two modal operators)

Normal modal logic: Syntax

Syntax of our language given by defining what its formulae are:

- Let Prop $= \{p, q, \dots\}$ be a countable set of variables for atomic propositions.
- Each $p \in$ Prop is a formula.
- $\top$ and $\bot$ are formulae.
- If φ, ψ are formulae, then so are $\neg\varphi$, $(\varphi \land \psi)$, $(\varphi \lor \psi)$, $(\varphi \to \psi)$, and $(\varphi \leftrightarrow \psi)$.
  (with the usual meaning as in ordinary propositional logic).
- If φ is a formula, then so are $\Box\varphi$ and $\Diamond\varphi$.

Normal modal logic: Semantics

Definition (Kripke model)

A **Kripke model** is a triple $M = \langle W, R, \pi \rangle$ where:

- W is a non-empty set of **possible worlds**,
- $R \subseteq W \times W$ is a binary relation on W (**accessibility relation**),
- π is a **valuation function** $\pi \colon W \to 2^{\mathsf{Prop}}$.

- R describes which worlds are considered possible relative to other worlds
- π specifies which atomic propositions are true in which world
- The pair $F = \langle W, R \rangle$ is called a **Kripke frame**.

Normal modal logic: Semantics

Satisfiability relation $\models$ between pairs (M, w) and formulae of the language is used to define semantics:

$$
\begin{array}{lcl}
(M,w) \models \top & & (M,w) \not\models \bot \\
(M,w) \models p & \iff & p \in \pi(w) \\
(M,w) \models \neg\varphi & \iff & (M,w) \not\models \varphi \\
(M,w) \models \varphi \rightarrow \psi & \iff & (M,w) \not\models \varphi \text{ or } (M,w) \models \psi \\
 & \ldots & \\
(M,w) \models \Box\varphi & \iff & (M,w') \models \varphi \text{ for each } w' \text{ s.t. } (w,w') \in R \\
(M,w) \models \Diamond\varphi & \iff & (M,w') \models \varphi \text{ for some } w' \text{ s.t. } (w,w') \in R
\end{array}
$$

Modal operators are duals of each other (like $\exists/\forall$):

$$\Box\varphi \leftrightarrow \neg\Diamond\neg\varphi$$

Satisfiability & validity concepts

A formula is called:

- **satisfiable** if it is satisfied for some (model, world) pair
- **unsatisfiable** if it is not satisfied for any (model, world) pair
- **true in a model** if it is satisfied for every world in the model
- **true in a frame** if it is satisfied for every world in each model based on the frame
- **valid in a class of models/frames** if it is true in every model/frame in the class
- **valid** if it is true in the class of all models (symb. $\models \varphi$)

Two basic properties:
K-axiom: $\models \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ is a valid formula
Necessitation rule: If $\models \varphi$, then $\models \Box\varphi$

Correspondence theory

- A **modal logic** is a set K of formulae (usually formulae valid in some class of frames)
- A member φ of K is called a **theorem** of the logic (denoted by: $\vdash_K \varphi$)
- Different sets of axioms correspond to different properties of the accessibility relation R (**correspondence theory**)
- Axioms are **characteristic** for a class of frames if they are valid in all and only those frames
- $\mathbf{K}\Sigma_1 \ldots \Sigma_n$ refers to the smallest modal logic containing axioms $\Sigma_1 \ldots \ \Sigma_n$

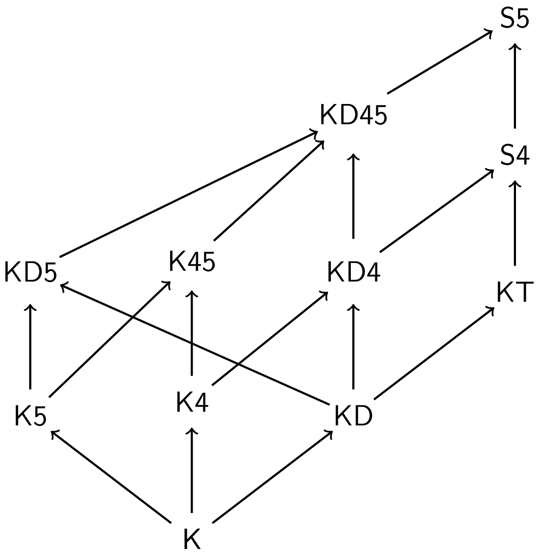
Correspondence theory

Correspondence between properties of R and axioms:

Name	Axiom	R	Characterization
T	$\Box\varphi \rightarrow \varphi$	reflexive	$\forall w.\ (w, w) \in R$
D	$\Box\varphi \rightarrow \Diamond\varphi$	serial	$\forall w \exists w'.\ (w, w') \in R$
4	$\Box\varphi \rightarrow \Box\Box\varphi$	transitive	$\forall w w' w''.\ (w, w') \in R \wedge (w', w'') \in R \Rightarrow (w, w'') \in R$
5	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$	Euclidean	$\forall w w' w''.\ (w, w') \in R \wedge (w, w'') \in R \Rightarrow (w', w'') \in R$

Some abbreviations often used: **KT** is called **T**, **KT4** is called **S4**, **KD45** is weak-S5, **KT5** called **S5**.

A lattice of modal logics



15.3 Epistemic logic

Normal modal logics as epistemic logics

- We assume: the agent knows a proposition A if A is true in all possible worlds that are accessible to the agent
- ... we use $\Box\varphi$ to represent "the agent knows that φ"
- In the case of several agents, models have to be extended to structures

$$\langle W, R_1, \ldots, R_n, \pi \rangle$$

  where R_i is the accessibility relation of agent i
- The single modal operator $\Box$ is replaced by unary modal operators K_i, one for each agent
- Semantically, replace the rule for $\Box$ by:

$$(M, w) \models \mathrm{K}_i \, \varphi \iff (M, w') \models \varphi \text{ for all } w' \text{ s.t. } (w, w') \in R_i$$

- We consider now multi-modal logics (e.g. $\mathbf{S5}_n$ multi-modal variant of **S5**)

Normal modal logics as epistemic logics

How well-suited are the properties of normal modal logic for describing knowledge and belief?

- Necessitation rule means that agents know all valid formulae (in particular all propositional logic tautologies)
- So agents always have an infinite amount of knowledge (sounds counterintuitive, but is it?)
- K-axiom causes a similar problem
  - Suppose φ is logical consequence of $\{\varphi_1, \dots, \varphi_n\}$
  - φ is true in every world in which $\varphi_1, \dots, \varphi_n$ are
  - Therefore $\varphi_1 \wedge \dots \wedge \varphi_n \rightarrow \varphi$ is valid
  - By necessitation, this rule must be believed
- By the K-axiom, the agent's knowledge is closed under logical consequence (if agent believes premises, it believes consequence)
- Agents know everything they might be able to infer!

Logical omniscience

Logical omniscience problem: all valid formulae are known and knowledge/belief is closed under logical consequence

- One problem concerns **consistency**:
  - human reasoners often have beliefs φ and ψ with $\varphi \vdash \neg\psi$ without being aware of inconsistency
  - thus ideal reasoners would believe **every** formula
- Second problem concerns **logical equivalence**:
  - Consider the following proposition:
    1. Hamlet's favourite colour is black
    2. Hamlet's favourite colour is black and every planar map can be four coloured
  - (2.) will be believed if and only if (1.) is believed, i.e. they are logically equivalent

Epistemic logic does not describe **actual beliefs** (mental states), but **rationality principles** for beliefs (cp. game theory, subjective probabilitites, . . .)

Axioms for knowledge and belief

How appropriate are the axioms D, T, 4, and 5 for logics of knowledge and belief?

- Axiom D requires that beliefs are not contradictory (reasonable): $K_i\,\varphi \rightarrow \neg\, K_i\,\neg\varphi$
- Axiom T (**truth/knowledge axiom**, requires that everything that is known is true: $K_i\,\varphi \rightarrow \varphi$

This can be used to distinguish knowledge from belief such that

"*i* **knows** φ **(if and) only if** *i* **believes** φ and φ is **true**"

Axioms for knowledge and belief

- Axiom 4 (**positive introspection**, requires that, if the agent knows a proposition, then s/he knows that s/he knows that: $\mathrm{K}_i\,\varphi \rightarrow \mathrm{K}_i\,\mathrm{K}_i\,\varphi$
- Axiom 5 (**negative introspection**, requires that, if the agent does not know a proposition, then s/he knows that s/he doesn't know that: $\neg\,\mathrm{K}_i\,\varphi \rightarrow \mathrm{K}_i\,\neg\,\mathrm{K}_i\,\varphi$
- Together axioms 4/5 can be read as: for every proposition φ, the agent knows **whether** s/he knows φ: $\mathrm{K}_i\,\mathrm{K}_i\,\varphi \vee \mathrm{K}_i\,\neg\,\mathrm{K}_i\,\varphi$
- Negative introspection considered more demanding than 4
- Usually, S5 is chosen as a logic of knowledge and KD45 as a knowledge of belief

Possible worlds in distributed systems

Consider the following simple model of a distributed system:

- **environment** described by a set E of environment states
- a set of n **processes**, each described by the same set L of local states
- the **system states** form a subset $G \subseteq E \times L \times \cdots \times L$, i.e., at each time point the system state is a tuple $(e, l_1, \ldots, l_n) \in G$
- **system** is specified by a set of **runs** of the system. i.e., maps $r \colon \mathbb{N} \to G$

Assume everything a process "knows" is encoded in the local state of the process. Thus, for runs r, r' and time points $t, t' \in \mathbb{N}$:

$$(r, t) \sim_i (r', t') \iff l_i = l'_i$$

with $r(t) = (e, \ldots, l_i, \ldots)$ and $r'(t') = (e', \ldots, l'_i, \ldots)$ — an equivalence relation. We can model what the processes know as S5-modalities.

Common and distributed knowledge

One would also like to model **common knowledge**, i.e. the propositions everyone knows, everyone knows that everyone knows, etc.

- Introduce first an operator for "**Everyone knows** that φ":

$$\mathsf{E}\,\varphi := \mathsf{K}_1\,\varphi \wedge \cdots \wedge \mathsf{K}_n\,\varphi$$

- Define then:

$$\mathsf{E}^1\,\varphi := \mathsf{E}\,\varphi \quad \text{and} \quad \mathsf{E}^{k+1}\,\varphi := \mathsf{E}\,\mathsf{E}^k\,\varphi$$

- Finally define the operator $\mathsf{C}\,\varphi$ for "**It is commonly known** that φ":

$$\mathsf{C}\,\varphi := \mathsf{E}\,\varphi \wedge \mathsf{E}^2\,\varphi \wedge \ldots$$

Infinite conjunction is quite a strong requirement, does common knowledge in this sense occur in practice?

Example

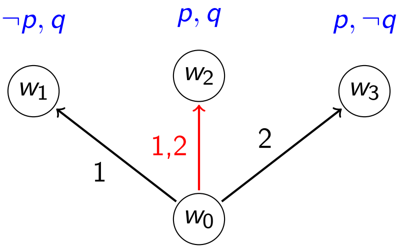
- Coordinated attack problem: two divisions of an army are camped on two hilltops waiting to attack enemy in the valley
- They can only attack successfully if they both attack at the same time
- Divisions can only communicate through messengers, communication takes time and may fail
- Even if messenger reaches other camp (e.g. with message "attack at dawn") generals can never be sure the message was received
- Awaiting confirmation does not solve problem, confirming party will never know whether other party received confirmation
- It turns out that no amount of communication is sufficient to bring about common knowledge

Common and distributed knowledge

- Another associated problem: distributed, implicit knowledge
- Assume an agent could read all other agents' minds this agent could have more knowledge than any other individual agent
- Example: one agent knows φ, the other (only) $\varphi \rightarrow \psi$, omniscient observer could infer ψ
- Distributed knowledge operator D can be introduced:

$$(M, w) \models \mathrm{D}\,\varphi \iff (M, w') \models \varphi, \text{ for each } w' \text{ with } (w, w') \in \bigcap_{i=1}^{n} R_i$$

Common and distributed knowledge



- Note that use of intersection rather than union actually increases knowledge
- These operators form a hierarchy:

$$\mathsf{C}\,\varphi \rightarrow \mathsf{E}^k\,\varphi \rightarrow \cdots \rightarrow \mathsf{E}\,\varphi \rightarrow \mathsf{K}_i\,\varphi \rightarrow \mathsf{D}\,\varphi$$

Example: Wise men puzzle

Situation: There are three wise men, and it is common knowledge that there are three red hats and two white hats. The king puts a hat on each of the wise man, and asks them sequentially whether they know the color of the hat on their head.

The first answers that he does not know. The second answers that he does not know. What does the third wise men answer?

Assumption: It is common knowledge that all wise men can see the hats on the others' heads, but not on their own. It is also common knowledge that every wise man hears the answers of the others.

15.4 Summary

- Thanks

Discussion

Are these logical models of practical use?

- Clearly, valuable for system specification (model checking)
- Inference intractable in most of these complex logics (e.g. satisfiability checking in S5 and KD45 is NP-complete, in K, S4, $S4_n$, $S5_n$ it is PSPACE-complete, with C most logics become EXPTIME-complete)
- Modal logic doesn't tell us anything about reasoning capabilities of agents themselves

Summary

- Logics for multiagent systems
- Logical modelling of mental states
- Modal logic as a popular method for doing that
- Possible-world semantics, correspondence theory
- Normal modal logics as epistemic logics
- Logical omniscience problems, critique
- Epistemic logic: common knowledge, distributed knowledge

Acknowledgments

These lecture slides are based on the following resources:

- Dr. Michael Rovatsos, The University of Edinburgh
  http://www.inf.ed.ac.uk/teaching/courses/abs/abs-timetable.html
- Michael Wooldridge: **An Introduction to MultiAgent Systems**, John Wiley & Sons, 2nd edition 2009.
- Ronald Fagin, J.Y. Halpern, Y. Moses, & M.Y. Vardi: **Reasoning about Knowledge**, MIT Press, 1995.