# Modal Logics

Knowledge Representation and Reasoning

November 7, 2005

# Outline

# Motivation for Studying Modal Logics

- Notions like believing and knowing require a more general semantics than e.g. propositional logic has.
- Some KR formalisms can be understood as (fragments of) a propositional modal logic.
- Application 1: spatial representation formalism RCC8.
- Application 2: description logics

# Motivation for Modal Logics

Often, we want to state something where we have an "embedded proposition":

- John believes that it is Sunday.
- I know that $2^{10} = 1024$.

Reasoning with embedded propositions:

- John believes that if it is Sunday then shops are closed.
- John believes that it is Sunday.
- This implies (assuming *belief* is closed under *modus ponens*):
  - John believes that shops are closed.

$\rightsquigarrow$ How to formalize this?

# Syntax

Propositional logic $+$ operators $\Box$ & $\Diamond$ (*Box & Diamond*):

$$\begin{array}{llll} \varphi & \longrightarrow & \ldots & \textit{classical propositional formula} \\ & \mid & \Box\varphi' & \textit{Box} \\ & \mid & \Diamond\varphi' & \textit{Diamond} \end{array}$$

$\Box$ and $\Diamond$ have the same operator precedence as $\neg$.

Some possible readings of $\Box\varphi$:

- Necessarily $\varphi$ (alethic)
- Always $\varphi$ (temporal)
- $\varphi$ should be true (deontic)
- Agent $A$ believes $\varphi$ (doxastic)
- Agent $A$ knows $\varphi$ (epistemic)

$\rightsquigarrow$ different semantics for different intended readings

# Truth Functional Semantics?

- Could it be possible to define the meaning of $\Box\varphi$ truth functionally, i.e. by referring to the truth value of $\varphi$ only?
- An attempt to interpret *necessity* truth-functionally:
  - If $\varphi$ is false, then $\Box\varphi$ should be false.
  - If $\varphi$ is true, then ...
    - ... $\Box\varphi$ should be true $\leadsto$ $\Box$ is the identity function
    - ... $\Box\varphi$ should be false $\leadsto$ $\Box\varphi$ is identical to falsity
- Note: There are only 4 different unary Boolean functions $\{T,F\} \to \{T,F\}$.

# Semantics: The Idea

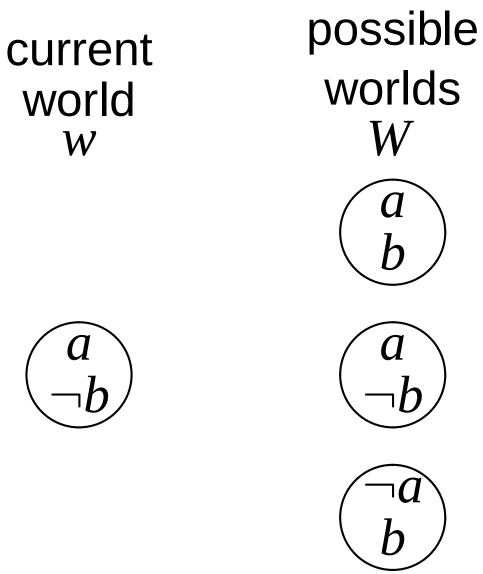
In classical propositional logic, formulae are interpreted over single interpretations and are evaluated to *true* or *false*.

In modal logics one considers sets of interpretations: possible worlds (physically possible, conceivable, . . . ).

Main idea:

- Consider a world (interpretation) $w$ and a *set of worlds* $W$ which are possible with respect to $w$.
- A classical formula (with no modal operators) $\varphi$ is true with respect to $(w,W)$ iff $\varphi$ is true in $w$.
- $\Box\varphi$ is true wrt $(w,W)$ iff $\varphi$ is true in all worlds in $W$.
- $\Diamond\varphi$ is true wrt $(w,W)$ iff $\varphi$ is true in one world in $W$.
- Meanings of $\Box$ and $\Diamond$ are interrelated by $\Diamond\varphi \equiv \neg\Box\neg\varphi$.

# Semantics: An Example




Examples:

- $a \wedge \neg b$ is true relative to $(w, W)$.
- $\Box a$ is not true relative to $(w, W)$.
- $\Box(a \vee b)$ is true relative to $(w, W)$.

Question: How to evaluate **modal** formulae in $w \in W$?

$\rightsquigarrow$ For each world, we specify a set of possible worlds.

$\rightsquigarrow$ frames

# Frames, Interpretations, and Worlds

A frame is a pair $\mathcal{F} = \langle W, R \rangle$, where $W$ is a non-empty set (of *worlds*) and $R \subseteq W \times W$ (the *accessibility relation*).
For $(w, v) \in R$ we write also $wRv$.
We say that $v$ is an $R$-successor of $w$ and that $v$ is reachable (or $R$-reachable) from $w$.
A ($\Sigma$)-interpretation (or model) based on the frame $\mathcal{F} = \langle W, R \rangle$ is a triple $\mathcal{I} = \langle W, R, \pi \rangle$, where $\pi$ is a function from worlds to truth assignments:

$$\pi \colon W \to (\Sigma \to \{T, F\})$$

# Semantics: Truth in one World

A formula $\varphi$ is true in world $w$ of an interpretation $I = \langle W, R, \pi \rangle$ under the following conditions:

$$
\begin{array}{rll}
I, w \models a & \textit{iff} & \pi(w)(a) = T \\
I, w \models \top & & \\
I, w \not\models \bot & & \\
I, w \models \neg\varphi & \textit{iff} & I, w \not\models \varphi \\
I, w \models \varphi \wedge \psi & \textit{iff} & I, w \models \varphi \text{ and } I, w \models \psi \\
I, w \models \varphi \vee \psi & \textit{iff} & I, w \models \varphi \text{ or } I, w \models \psi \\
I, w \models \varphi \rightarrow \psi & \textit{iff} & \text{if } I, w \models \varphi \text{ then } I, w \models \psi \\
I, w \models \varphi \leftrightarrow \psi & \textit{iff} & I, w \models \varphi \text{ if and only if } I, w \models \psi \\
I, w \models \Box\varphi & \textit{iff} & I, u \models \varphi \text{ for all } u \text{ s.t. } wRu \\
I, w \models \Diamond\varphi & \textit{iff} & I, u \models \varphi \text{ for at least one } u \text{ s.t. } wRu
\end{array}
$$

# Satisfiability and Validity

A formula $\varphi$ is satisfiable in an interpretation $\mathcal{I}$ (or *in a frame* $\mathcal{F}$, or *in a class of frames* $\mathcal{C}$) if there exists a world in $\mathcal{I}$ (or an interpretation $\mathcal{I}$ based on $\mathcal{F}$, or an interpretation $\mathcal{I}$ based on a frame contained in the class $\mathcal{C}$, respectively) such that $\mathcal{I}, w \models \varphi$.

A formula $\varphi$ is true in an interpretation $\mathcal{I}$ (symbolically $\mathcal{I} \models \varphi$) if $\varphi$ is true in all worlds of $\mathcal{I}$.

A formula $\varphi$ is valid in a frame $\mathcal{F}$ or $\mathcal{F}$ *-valid* (symbolically $\mathcal{F} \models \varphi$) if $\varphi$ is true in all interpretations based on $\mathcal{F}$.

A formula $\varphi$ is valid in a class of frames $\mathcal{C}$ or $\mathcal{C}$*-valid* (symbolically $\mathcal{C} \models \varphi$) if $\mathcal{F} \models \varphi$ for all $\mathcal{F} \in \mathcal{C}$.

K is the class of all frames – named after *Saul Kripke*, who invented this semantics.

# Validity: Some Examples

1. $\varphi \vee \neg\varphi$
2. $\Box(\varphi \vee \neg\varphi)$
3. $\Box\varphi$, if $\varphi$ is a classical tautology
4. $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ (*axiom schema $K$*)

# Validity: Some Examples

## Theorem

*$K$ is **K**-valid.* *($K$ = $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$)*

## Proof.

Let $\mathcal{I}$ be an interpretation and let $w$ be a world in $\mathcal{I}$.
Assumption: $\mathcal{I}, w \models \Box(\varphi \rightarrow \psi)$, i.e., in all worlds $u$ with $wRu$, if $\varphi$ is true then also $\psi$ is. (Otherwise $K$ is true in any case.)
If $\Box\varphi$ is false in $w$, then $(\Box\varphi \rightarrow \Box\psi)$ is true and $K$ is true in $w$.
If $\Box\varphi$ is true in $w$, then both $\Box(\varphi \rightarrow \psi)$ and $\Box\varphi$ are true in $w$. Hence both $\varphi \rightarrow \psi$ and $\varphi$ are true in every world $u$ accessible from $w$. Hence $\psi$ is true in any such $u$, and therefore $w \models \Box\psi$. Since $\mathcal{I}$ and $w$ were arbitrary, the argument goes through for any $\mathcal{I}, w$, i.e., $K$ is **K**-valid. □

# Non-validity: Example

### Proposition

*$\Diamond\top$ is not **K**-valid.*

### Proof.

A counterexample is the following interpretation:

$$\mathcal{I} = \langle \{w\}, \emptyset, \{w \mapsto (a \mapsto T)\} \rangle.$$

We have $\mathcal{I}, w \not\models \Diamond\top$ because there is no $u$ such that $wRu$. □

# Non-validity: Example

## Proposition

*$\Box\varphi \rightarrow \varphi$ is not* **K***-valid.*

## Proof.

A counterexample is the following interpretation:

$$I = \langle \{w\}, \emptyset, \{w \mapsto (a \mapsto F)\} \rangle.$$

We have $I, w \models \Box a$ but $I, w \not\models a$. □

# Non-validity: Another Example

## Proposition

$\Box\varphi \rightarrow \Box\Box\varphi$ *is not* **K**-*valid.*

## Proof.

A counterexample is the following interpretation:

$$I = \langle \{u, v, w\}, \{(u, v), (v, w)\}, \pi \rangle$$

with

$$\begin{array}{rcl} \pi(u) & = & \{a \mapsto T\} \\ \pi(v) & = & \{a \mapsto T\} \\ \pi(w) & = & \{a \mapsto F\} \end{array}$$

This means $I, u \models \Box a$, but $I, u \not\models \Box\Box a$. □

# Accessibility and Axiom Schemata

Let us consider the following axiom schemata:

- T: $\Box\varphi \rightarrow \varphi$ (*knowledge axiom*)
- 4: $\Box\varphi \rightarrow \Box\Box\varphi$ (*positive introspection*)
- 5: $\Diamond\varphi \rightarrow \Box\Diamond\varphi$ (or $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$: *negative introspection*)
- B: $\varphi \rightarrow \Box\Diamond\varphi$
- D: $\Box\varphi \rightarrow \Diamond\varphi$ (or $\Box\varphi \rightarrow \neg\Box\neg\varphi$: *disbelief in the negation*)

. . . and the following classes of frames, for which the accessibility relation is restricted as follows:

- T: reflexive ($wRw$ for each world $w$)
- 4: transitive ($wRu$ and $uRv$ implies $wRv$)
- 5: euclidian ($wRu$ and $wRv$ implies $uRv$)
- B: symmetric ($wRu$ implies $uRw$)
- D: serial (for each $w$ there exists $v$ with $wRv$)

# Connection between Accessibility Relations and Axiom Schemata (1)

## Theorem

*Axiom schema $T$ $(4, 5, B, D)$ is* **T**- *valid* (**4-, 5-, B-**, *or* **D**-*valid, respectively).*

## Proof.

For $T$ and **T**:Let $\mathcal{F}$ be a frame from class **T**. Let $\mathcal{I}$ be an interpretation based on $\mathcal{F}$ and let $w$ be an arbitrary world in $\mathcal{I}$. If $\Box\varphi$ is not true in a world $w$, then axiom $T$ is true in $w$. If $\Box\varphi$ is true in $w$, then $\varphi$ is true in all accessible worlds.Since the accessibility relation is reflexive, $w$ is among the accessible worlds, i.e., $\varphi$ is true in $w$. This means that also in this case $T$ is true $w$. This means, $T$ is true in all worlds in all interpretations based on **T**-frames. □

# Connection between Accessibility Relations and Axiom Schemata (2)

### Theorem

*If $T$ $(4, 5, B, D)$ is valid in a frame $\mathcal{F}$, then $\mathcal{F}$ is a* **T***-Frame* **(4-, 5-, B-***, or* **D***-frame, respectively).*

### Proof.

For $T$ and **T**: Assume that $\mathcal{F}$ is not a **T**-frame. We will construct an interpretation based on $\mathcal{F}$ that falsifies $T$.
Because $\mathcal{F}$ is not a **T**-frame, there is a world $w$ such that not $wRw$.
Construct an interpretation $\mathcal{I}$ such that $w \not\models p$ and $v \models p$ for all $v$ such that $wRv$.
Now $w \models \Box p$ and $w \not\models p$, and hence $w \not\models \Box p \rightarrow p$. □

# Different Modal Logics

| Name | Property | Axiom schema |
|---|---|---|
| $K$ | – | $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ |
| $T$ | reflexivity | $\Box\varphi \to \varphi$ |
| 4 | transitivity | $\Box\varphi \to \Box\Box\varphi$ |
| 5 | euclidicity | $\Diamond\varphi \to \Box\Diamond\varphi$ |
| $B$ | symmetry | $\varphi \to \Box\Diamond\varphi$ |
| $D$ | seriality | $\Box\varphi \to \Diamond\varphi$ |

Some basic modal logics:

$$\begin{array}{rcl} K & & \\ KT4 & = & S4 \\ KT5 & = & S5 \\ \vdots & & \end{array}$$

# Different Modal Logics

| logics | $\Box$ | $\Diamond = \neg\Box\neg$ | K | T | 4 | 5 | B | D |
|---|---|---|---|---|---|---|---|---|
| alethic | necessarily | possibly | Y | Y | Y | Y | Y | Y |
| epistemic | known | possible | Y | Y | Y | Y | Y | Y |
| doxastic | believed | possible | Y | N | Y | Y | N | Y |
| deontic | obligatory | permitted | Y | N | ?/Y | ?/Y | N | Y |
| temporal | always in the future | sometimes | Y | Y/N | Y | N | N | N/Y |

# Proof Methods

- How can we show that a formula is $\mathcal{C}$-valid?
- In order to show that a formula is not $\mathcal{C}$-valid, one can construct a counterexample (= an interpretation that falsifies it.)
- When trying out all ways of generating a counterexample without success, this counts as a proof of validity.
- method of (analytic/semantic) tableaux

# Tableau Method

A tableau is a tree with nodes marked as follows:

- $w \models \varphi$,
- $w \not\models \varphi$, and
- $wRv$.

A branch that contains nodes marked with $w \models \varphi$ and $w \not\models \varphi$ is closed. All other branches are open. If all branches are closed, the tableau is closed.

A tableau is constructed by using the tableau rules.

# Tableau Rules for the Propositional Logic

$$\frac{w \models \varphi \vee \psi}{w \models \varphi \mid w \models \psi}$$

$$\frac{w \not\models \varphi \vee \psi}{\begin{array}{c} w \not\models \varphi \\ w \not\models \psi \end{array}}$$

$$\frac{w \models \neg\varphi}{w \not\models \varphi}$$

$$\frac{w \models \varphi \wedge \psi}{\begin{array}{c} w \models \varphi \\ w \models \psi \end{array}}$$

$$\frac{w \not\models \varphi \wedge \psi}{w \not\models \varphi \mid w \not\models \psi}$$

$$\frac{w \not\models \neg\varphi}{w \models \varphi}$$

$$\frac{w \models \varphi \rightarrow \psi}{w \not\models \varphi \mid w \models \psi}$$

$$\frac{w \not\models \varphi \rightarrow \psi}{\begin{array}{c} w \models \varphi \\ w \not\models \psi \end{array}}$$

# Additional Tableau Rules for the Modal Logic **K**

$$\frac{w \models \Box\varphi}{v \models \varphi}$$ if $wRv$ is on the branch already

$$\frac{w \not\models \Box\varphi}{\begin{array}{c} wRv \\ v \not\models \varphi \end{array}}$$ for new $v$

$$\frac{w \models \Diamond\varphi}{\begin{array}{c} wRv \\ v \models \varphi \end{array}}$$ for new $v$

$$\frac{w \not\models \Diamond\varphi}{v \not\models \varphi}$$ if $wRv$ is on the branch already

# Properties of **K** Tableaux

### Proposition

*If a $K$-tableau is closed, the truth condition at the root cannot be satisfied.*

### Theorem (Soundness)

*If a $K$-tableau with root $w \not\models \varphi$ is closed, then $\varphi$ is **K**-valid.*

### Theorem (Completeness)

*If $\varphi$ is **K**-valid, then there is a closed tableau with root $w \not\models \varphi$.*

### Proposition (Termination)

*There are strategies for constructing **K**-tableaux that always terminate after a finite number of steps, and result in a closed tableau whenever one exists.*

# Tableau Rules for Other Modal Logics

Proofs within more restricted classes of frames allow the use of further tableau rules.

- For reflexive (**T**) frames we may extend any branch with $wRw$.
- For transitive (**4**) frames we have the following additional rule:
  - If $wRv$ and $vRu$ are in a branch, $wRu$ may be added to the branch.
- For serial (**D**) frames we have the following rule:
  - If there is $w \models \ldots$ or $w \not\models \ldots$ on a branch, then add $wRv$ for a new world $v$.
- Similar rules for other properties...

# Testing Logical Consequence with Tableaux

- Let $\Theta$ be a set of formulas. When does a formula $\varphi$ follow from $\Theta$: $\Theta \models_{\mathbf{X}} \varphi$?
- Test whether in all interpretations on $\mathbf{X}$-frames in which $\Theta$ is true, also $\varphi$ is true.
- Wouldn't there be a deduction theorem we could use?
- Example: $a \models_{\mathbf{K}} \Box a$ holds, but $a \rightarrow \Box a$ is not **K**-valid.
- There is no deduction theorem as in the propositional logic, and logical consequence cannot be directly reduced to validity!

# Tableaus and Logical Implication

For testing logical consequence, we can use the following tableau rule:

- If $w$ is a world on a branch and $\psi \in \Theta$, then we can add $w \models \psi$ to our branch.
- Soundness is obvious.
- Completeness is non-trivial.

# Connection between propositional modal logic and FOL?

- There are similarities between the predicate logic and propositional modal logics:
    1. $\Box$ vs. $\forall$
    2. $\Diamond$ vs. $\exists$
    3. the possible worlds vs. the objects of the universe
- In fact, we can show for many propositional modal logics that they can be embedded in the predicate logic. $\Longrightarrow$ Modal logics can be understood as a sublanguage of FOL.

# Embedding Modal Logics in the Predicate Logic (1)

1. $\tau(p, x) = p(x)$ for propositional variables $p$
2. $\tau(\neg\phi, x) = \neg\tau(\phi, x)$
3. $\tau(\phi \vee \psi, x) = \tau(\phi, x) \vee \tau(\psi, x)$
4. $\tau(\phi \wedge \psi, x) = \tau(\phi, x) \wedge \tau(\psi, x)$
5. $\tau(\Box\phi, x) = \forall y(R(x, y) \rightarrow \tau(\phi, y))$ for some new $y$
6. $\tau(\Diamond\phi, x) = \exists y(R(x, y) \wedge \tau(\phi, y))$ for some new $y$

# Embedding Modal Logics in the Predicate Logic (2)

### Theorem

$\phi$ *is K-valid if and only if* $\forall x \tau(\phi, x)$ *is valid in the predicate logic.*

### Theorem

$\phi$ *is T-valid if and only if in the predicate logic the logical consequence* $\{\forall x R(x,x)\} \models \forall x \tau(\phi, x)$ *holds.*

### Example

$((\Box p) \wedge \Diamond(p \to q)) \to \Diamond q$ is K-valid because

$$\forall x((\forall x'(R(x,x') \to p(x'))) \wedge \exists x'(R(x,x') \wedge (p(x') \to q(x')))) \\ \to \exists x'(R(x,x') \wedge q(x'))$$

is valid in the predicate logic.

# Outlook

We only looked at some basic propositional modal logics. There are also

- modal first order logics (with quantification $\forall$ and $\exists$ and predicates)
- multi-modal logics: more than one modality, e.g. knowledge/belief operators for several agents
- temporal and dynamic logics (modalities that refer to time or programs, respectively)

# Outlook

Did we really do something new? Couldn't we have done everything in propositional modal logic in FOL already?

- Yes – but now we know much more about the (restricted) system and have decidable problems!

# Literature


Anil Nerode.
Some lectures on modal logic.
In F. L. Bauer, editor, *Logic, Algebra, and Computation*, volume 79 of *NATO ASI Series on Computer and System Sciences*, pages 281–334. Springer-Verlag, Berlin, Heidelberg, New York, 1991.

Melvin Fitting.
Basic Modal Logic.
In D. M. Gabbay and C. J. Hogger and J. A. Robinson, eds., *Handbook of Logic in Artificial Intelligence and Logic Programming – Vol. 1: Logical Foundations*, Oxford University Press, Oxford, UK, 1993.

M. Fitting.
*Proof Methods for Modal and Intuitionistic Logic*.
Reidel, 1983.

Robert Goldblatt.
*Logics of Time and Computation*.
Stanford University, 1992.

B. F. Chellas.
*Modal Logic: An Introduction*.
Cambridge University, 1980.