09—Modal Logic II

CS 3234: Logic and Formal Systems

Martin Henz and Aquinas Hobor

October 14, 2010

Generated on Thursday 14th October, 2010, 11:39
Review of Modal Logic

Some Modal Logics

Natural Deduction in Modal Logic

Knowledge in Multi-Agent Systems
1. Review of Modal Logic
   - Motivation
   - Syntax and Semantics
   - Valid Formulas wrt Modalities
   - Correspondence Theory

2. Some Modal Logics

3. Natural Deduction in Modal Logic

4. Knowledge in Multi-Agent Systems
Notions of Truth

- Often, it is not enough to distinguish between “true” and “false”.
- We need to consider modalities if truth, such as:
  - necessity
  - time
  - knowledge by an agent
- Modal logic constructs a framework using which modalities can be formalized and reasoning methods can be established.
Syntax of Basic Modal Logic

\[ \phi ::= \top | \bot | p | (\neg \phi) | (\phi \land \phi) \\
| (\phi \lor \phi) | (\phi \to \phi) \\
| (\phi \leftrightarrow \phi) \\
| (\Box \phi) | (\Diamond \phi) \]
Definition

A model $\mathcal{M}$ of propositional modal logic over a set of propositional atoms $A$ is specified by three things:

1. A $W$ of worlds;
Definition

A model $\mathcal{M}$ of propositional modal logic over a set of propositional atoms $A$ is specified by three things:

1. A $W$ of worlds;
2. a relation $R$ on $W$, meaning $R \subseteq W \times W$, called the accessibility relation;
Definition

A model $M$ of propositional modal logic over a set of propositional atoms $A$ is specified by three things:

1. A $W$ of worlds;
2. a relation $R$ on $W$, meaning $R \subseteq W \times W$, called the accessibility relation;
3. a function $L : W \rightarrow A \rightarrow \{T, F\}$, called labeling function.
Example

\[ W = \{x_1, x_2, x_3, x_4, x_5, x_6\} \]
\[ R = \{(x_1, x_2), (x_1, x_3), (x_2, x_2), (x_2, x_3), (x_3, x_2), (x_4, x_5), (x_5, x_4), (x_5, x_6)\} \]
\[ L = \{(x_1, \{q\}), (x_2, \{p, q\}), (x_3, \{p\}), (x_4, \{q\}), (x_5, \{}), (x_6, \{p\})\} \]
When is a formula true in a possible world?

**Definition**

Let $\mathcal{M} = (W, R, L)$, $x \in W$, and $\phi$ a formula in basic modal logic. We define $x \models \phi$ via structural induction:

- $x \models \top$
- $x \not\models \bot$
- $x \models p$ iff $p \in L(x)(p) = T$
- $x \models \neg \phi$ iff $x \not\models \phi$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
- $x \models \phi \lor \psi$ iff $x \models \phi$ or $x \models \psi$
- ...


When is a formula true in a possible world?

Definition (continued)
Let $\mathcal{M} = (W, R, L)$, $x \in W$, and $\phi$ a formula in basic modal logic. We define $x \vDash \phi$ via structural induction:

- ...  
- $x \vDash \phi \to \psi$ iff $x \vDash \psi$, whenever $x \vDash \phi$
- $x \vDash \phi \iff \psi$ iff ($x \vDash \phi$ iff $x \vDash \psi$)
- $x \vDash \lozenge \phi$ iff for each $y \in W$ with $R(x, y)$, we have $y \vDash \phi$
- $x \vDash \Box \phi$ iff there is a $y \in W$ such that $R(x, y)$ and $y \vDash \phi$. 
Example

- $x_1 \models q$
- $x_1 \models \diamond q$, $x_1 \not\models \Box q$
- $x_5 \not\models \Box p$, $x_5 \not\models \Box q$, $x_5 \not\models \Box (p \lor q)$, $x_5 \models \Box (p \lor q)$
- $x_6 \models \Box \phi$ holds for all $\phi$, but $x_6 \not\models \diamond \phi$
A Range of Modalities

In a particular context \( \square \phi \) could mean:

- It is necessarily true that \( \phi \)
- It will always be true that \( \phi \)
- It ought to be that \( \phi \)
- Agent \( Q \) believes that \( \phi \)
- Agent \( Q \) knows that \( \phi \)
- After any execution of program \( P \), \( \phi \) holds.

Since \( \Diamond \phi \equiv \neg \square \neg \phi \), we can infer the meaning of \( \Diamond \) in each context.
### A Range of Modalities

From the meaning of $\Box \phi$, we can conclude the meaning of $\Diamond \phi$, since $\Diamond \phi \equiv \neg \Box \neg \phi$:

<table>
<thead>
<tr>
<th>$\Box \phi$</th>
<th>$\Diamond \phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is necessarily true that $\phi$</td>
<td>It is possibly true that $\phi$</td>
</tr>
<tr>
<td>It will always be true that $\phi$</td>
<td>Sometime in the future $\phi$</td>
</tr>
<tr>
<td>It ought to be that $\phi$</td>
<td>It is permitted to be that $\phi$</td>
</tr>
<tr>
<td>Agent $Q$ believes that $\phi$</td>
<td>$\phi$ is consistent with $Q$’s beliefs</td>
</tr>
<tr>
<td>Agent $Q$ knows that $\phi$</td>
<td>For all $Q$ knows, $\phi$</td>
</tr>
<tr>
<td>After any run of $P$, $\phi$ holds.</td>
<td>After some run of $P$, $\phi$ holds</td>
</tr>
</tbody>
</table>
### Formula Schemes that hold wrt some Modalities

<table>
<thead>
<tr>
<th>□ φ</th>
<th>□ φ → φ</th>
<th>□ φ → □ φ</th>
<th>□ φ → □□ φ</th>
<th>□ φ → □ φ → □ φ</th>
<th>□ φ → □ φ → □ φ</th>
<th>□ φ → □ φ → □ φ</th>
<th>□ (φ → ψ) ∧ □ φ</th>
<th>□ (φ ∧ ψ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

- **It is necessary that φ**
- **It will always be that φ**
- **It ought to be that φ**
- **Agent Q believes that φ**
- **Agent Q knows that φ**
- **After running P, φ**

The table shows validity for different modal formulas and interpretations.
## Modalities lead to Interpretations of $R$

<table>
<thead>
<tr>
<th>$\Box \phi$</th>
<th>$R(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is necessarily true that $\phi$</td>
<td>$y$ is possible world according to info at $x$</td>
</tr>
<tr>
<td>It will always be true that $\phi$</td>
<td>$y$ is a future world of $x$</td>
</tr>
<tr>
<td>It ought to be that $\phi$</td>
<td>$y$ is an acceptable world according to the information at $x$</td>
</tr>
<tr>
<td>Agent Q believes that $\phi$</td>
<td>$y$ could be the actual world according to Q’s beliefs at $x$</td>
</tr>
<tr>
<td>Agent Q knows that $\phi$</td>
<td>$y$ could be the actual world according to Q’s knowledge at $x$</td>
</tr>
<tr>
<td>After any execution of $P$, $\phi$ holds</td>
<td>$y$ is a possible resulting state after execution of $P$ at $x$</td>
</tr>
</tbody>
</table>
Possible Properties of $R$

- reflexive: for every $w \in W$, we have $R(x, x)$.
- symmetric: for every $x, y \in W$, we have $R(x, y)$ implies $R(y, x)$.
- serial: for every $x$ there is a $y$ such that $R(x, y)$.
- transitive: for every $x, y, z \in W$, we have $R(x, y)$ and $R(y, z)$ imply $R(x, z)$.
- Euclidean: for every $x, y, z \in W$ with $R(x, y)$ and $R(x, z)$, we have $R(y, z)$.
- functional: for each $x$ there is a unique $y$ such that $R(x, y)$.
- linear: for every $x, y, z \in W$ with $R(x, y)$ and $R(x, z)$, we have $R(y, z)$ or $y = z$ or $R(z, y)$.
- total: for every $x, y \in W$, we have $R(x, y)$ and $R(y, x)$.
- equivalence: reflexive, symmetric and transitive.
Theorem

The following statements are equivalent:

- $R$ is reflexive;
- $F$ satisfies $\square \phi \rightarrow \phi$;
- $F$ satisfies $\square p \rightarrow p$;

Theorem

The following statements are equivalent:

- $R$ is transitive;
- $F$ satisfies $\square \phi \rightarrow \square \square \phi$;
- $F$ satisfies $\square p \rightarrow \square \square p$;
### Formula Schemes and Properties of $R$

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>$\Box \phi \rightarrow \phi$</td>
<td>reflexive</td>
</tr>
<tr>
<td>B</td>
<td>$\phi \rightarrow \Box \Diamond \phi$</td>
<td>symmetric</td>
</tr>
<tr>
<td>D</td>
<td>$\Box \phi \rightarrow \Diamond \phi$</td>
<td>serial</td>
</tr>
<tr>
<td>4</td>
<td>$\Box \phi \rightarrow \Box \Box \phi$</td>
<td>transitive</td>
</tr>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
<tr>
<td></td>
<td>$\Box \phi \leftrightarrow \Diamond \phi$</td>
<td>functional</td>
</tr>
<tr>
<td></td>
<td>$\Box (\phi \land \Box \phi \rightarrow \psi) \lor \Box (\psi \land \Box \psi \rightarrow \phi)$</td>
<td>linear</td>
</tr>
</tbody>
</table>
1. Review of Modal Logic

2. Some Modal Logics
   - K
   - KT45
   - KT4

3. Natural Deduction in Modal Logic

4. Knowledge in Multi-Agent Systems
Which Formula Schemes to Choose?

Definition

Let $\mathcal{L}$ be a set of formula schemes and $\Gamma \cup \{\psi\}$ a set of formulas of basic modal logic.
Which Formula Schemes to Choose?

Definition

Let $\mathcal{L}$ be a set of formula schemes and $\Gamma \cup \{\psi\}$ a set of formulas of basic modal logic.

- A set of formula schemes is said to be closed iff it contains all substitution instances of its elements.
Which Formula Schemes to Choose?

**Definition**

Let $\mathcal{L}$ be a set of formula schemes and $\Gamma \cup \{\psi\}$ a set of formulas of basic modal logic.

- A set of formula schemes is said to be *closed* iff it contains all substitution instances of its elements.
- Let $\mathcal{L}_c$ be the smallest closed superset of $\mathcal{L}$. 
Which Formula Schemes to Choose?

Definition

Let $\mathcal{L}$ be a set of formula schemes and $\Gamma \cup \{\psi\}$ a set of formulas of basic modal logic.

- A set of formula schemes is said to be *closed* iff it contains all substitution instances of its elements.
- Let $\mathcal{L}_c$ be the smallest closed superset of $\mathcal{L}$.
- $\Gamma$ entails $\psi$ in $\mathcal{L}$ iff $\Gamma \cup \mathcal{L}_c$ semantically entails $\psi$. We say $\Gamma \models_{\mathcal{L}} \psi$. 


Examples of Modal Logics: $K$

$K$ is the weakest modal logic, $\mathcal{L} = \emptyset$. 
Examples of Modal Logics: KT45

\[ \mathcal{L} = \{ T, 4, 5 \} \]
Examples of Modal Logics: KT45

\[ \mathcal{L} = \{ T, 4, 5 \} \]

Used for reasoning about knowledge.
Examples of Modal Logics: KT45

\( \mathcal{L} = \{ T, 4, 5 \} \)

Used for reasoning about knowledge.

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of ( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>( \Box \phi \rightarrow \phi )</td>
<td>reflexive</td>
</tr>
<tr>
<td>4</td>
<td>( \Box \phi \rightarrow \Box \Box \phi )</td>
<td>transitive</td>
</tr>
<tr>
<td>5</td>
<td>( \Diamond \phi \rightarrow \Box \Diamond \phi )</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>
Examples of Modal Logics: KT45

\[ \mathcal{L} = \{ T, 4, 5 \} \]

Used for reasoning about knowledge.

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of ( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( \square \phi \rightarrow \phi )</td>
<td>reflexive</td>
</tr>
<tr>
<td>( 4 )</td>
<td>( \square \phi \rightarrow \square \square \phi )</td>
<td>transitive</td>
</tr>
<tr>
<td>( 5 )</td>
<td>( \Diamond \phi \rightarrow \square \Diamond \phi )</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

- **T**: Truth: agent \( Q \) only knows true things.
- **4**: Positive introspection: If \( Q \) knows something, he knows that he knows it.
- **5**: Negative introspection: If \( Q \) doesn’t know something, he knows that he doesn’t know it.
**Explanation of Negative Introspection**

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>
## Explanation of Negative Introspection

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

\[
\Diamond \phi \rightarrow \Box \Diamond \phi
\]
Explanation of Negative Introspection

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

$$
\begin{align*}
\Diamond \phi & \rightarrow \Box \Diamond \phi \\
\Diamond \neg \psi & \rightarrow \Box \Diamond \neg \psi
\end{align*}
$$
Explanation of Negative Introspection

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

$\Diamond \phi \rightarrow \Box \Diamond \phi$

$\Diamond \neg \psi \rightarrow \Box \Diamond \neg \psi$

$\neg \Box \neg \neg \psi \rightarrow \Box \neg \neg \neg \neg \neg \neg \psi$
### Explanation of Negative Introspection

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\Diamond \phi & \rightarrow \Box \Diamond \phi \\
\Diamond \neg \psi & \rightarrow \Box \Diamond \neg \psi \\
\neg \Box \neg \neg \psi & \rightarrow \Box \neg \Box \neg \neg \psi \\
\neg \Box \psi & \rightarrow \Box \neg \Box \psi
\end{align*}
\]
### Explanation of Negative Introspection

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of $R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$\Diamond \phi \rightarrow \Box \Diamond \phi$</td>
<td>Euclidean</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\Diamond \phi & \rightarrow \Box \Diamond \phi \\
\Diamond \neg \psi & \rightarrow \Box \Diamond \neg \psi \\
\neg \Box \neg \neg \psi & \rightarrow \Box \neg \Box \neg \neg \psi \\
\neg \Box \psi & \rightarrow \Box \neg \Box \psi
\end{align*}
\]

If $Q$ doesn’t know $\psi$, he knows that he doesn’t know $\psi$. 
Accessibility relations for KT45

KT45 hold if and only if $R$ is reflexive (T), transitive (4) and Euclidean (5).
Correspondence for KT45

Accessibility relations for KT45

KT45 hold if and only if $R$ is reflexive ($T$), transitive ($4$) and Euclidean ($5$).

Fact on such relations

A relation is reflexive, transitive and Euclidean iff it is reflexive, transitive and symmetric, i.e. iff it is an equivalence relation.
Collapsing Modalities

**Theorem**

Any sequence of modal operators and negations is KT45 is equivalent to one of the following: $\neg$, $\Box$, $\Diamond$, $\neg\neg\Box$, and $\neg\Diamond$, where $\neg$ indicates the absence of any negation or modality.
Collapsing Modalities

**Theorem**

Any sequence of modal operators and negations is KT45 is equivalent to one of the following: $\neg$, $\Box$, $\Diamond$, $\neg\Box$, and $\neg\Diamond$, where $\neg$ indicates the absence of any negation or modality.

**Examples**

1. $\Box\Box\phi \equiv \Box\phi$
Collapsing Modalities

Theorem

Any sequence of modal operators and negations is KT45 is equivalent to one of the following: \( \neg, \Box, \Diamond, \neg, \neg\Box, \) and \( \neg\Diamond, \) where \( \neg \) indicates the absence of any negation or modality.

Examples

- \( \Box\Box\phi \equiv \Box\phi \)
- \( \Diamond\Box\phi \equiv \Diamond\phi \)
Collapsing Modalities

**Theorem**

Any sequence of modal operators and negations is KT45 is equivalent to one of the following: \(\neg, \Box, \Diamond, \neg, \neg\Box, \) and \(\neg\Diamond,\) where \(\neg\) indicates the absence of any negation or modality.

**Examples**

- \(\Box\Box\phi \equiv \Box\phi\)
- \(\Diamond\Box\phi \equiv \Diamond\phi\)
- \(\neg\Diamond\neg\phi \equiv \Box\phi\)
Examples of Modal Logics: KT4

\[ \mathcal{L} = \{ T, 4 \} \]

Used for partial evaluation in computer science.
Examples of Modal Logics: KT4

\[ \mathcal{L} = \{ T, 4 \} \]

Used for partial evaluation in computer science.

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of ( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( \Box \phi \rightarrow \phi )</td>
<td>reflexive</td>
</tr>
<tr>
<td>( 4 )</td>
<td>( \Box \phi \rightarrow \Box \Box \phi )</td>
<td>transitive</td>
</tr>
</tbody>
</table>
Examples of Modal Logics: KT4

\[ \mathcal{L} = \{ T, 4 \} \]

Used for partial evaluation in computer science.

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of ( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( \Box \phi \rightarrow \phi )</td>
<td>reflexive</td>
</tr>
<tr>
<td>( 4 )</td>
<td>( \Box \phi \rightarrow \Box \Box \phi )</td>
<td>transitive</td>
</tr>
</tbody>
</table>

- **T**: Truth: agent \( Q \) only knows true things.
Examples of Modal Logics: KT4

\[ \mathcal{L} = \{ T, 4 \} \]

Used for partial evaluation in computer science.

<table>
<thead>
<tr>
<th>name</th>
<th>formula scheme</th>
<th>property of ( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( \square \phi \rightarrow \phi )</td>
<td>reflexive</td>
</tr>
<tr>
<td>( 4 )</td>
<td>( \square \phi \rightarrow \square \square \phi )</td>
<td>transitive</td>
</tr>
</tbody>
</table>

- \( T \): Truth: agent \( Q \) only knows true things.
- \( 4 \): Positive introspection: If \( Q \) knows something, he knows that he knows it.
Correspondence for KT4

Accessibility relations for KT4

KT4 hold if and only if \( R \) is reflexive (T), and transitive (4).
Correspondence for KT4

Accessibility relations for KT4

KT4 hold if and only if $R$ is reflexive (T), and transitive (4).

Definition

A reflexive and transitive relation is called a preorder.
Collapsing Modalities

Theorem

Any sequence of modal operators and negations is KT4 is equivalent to one of the following:

\(-, \square, \Diamond, \square\Diamond, \Diamond\square, \square\Diamond\square, \neg, \neg\square, \neg\Diamond, \neg\square\Diamond, \neg\Diamond\square, \neg\square\Diamond\square, \neg\Diamond\square\Diamond, \text{ and } \neg\Diamond\square\Diamond.\)
Definition

A model of intuitionistic propositional logic is a model \( M = (W, R, L) \) of KT4 such that \( R(x, y) \) always implies \( L(x)(p) \to L(y)(p) \).
Satisfaction in Intuitionistic Logic

Definition

We change the definition of $x \models \phi$ as follows:

- $x \models \top$
- $x \not\models \bot$
- $x \models p$ iff $p \in L(x)$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
- $x \models \phi \lor \psi$ iff $x \models \phi$ or $x \models \psi$

as usual,
We change the definition of $x \models \phi$ as follows:

- $x \models \top$
- $x \not\models \bot$
- $x \models p$ iff $p \in L(x)$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
- $x \models \phi \lor \psi$ iff $x \models \phi$ or $x \models \psi$

as usual, but now:

- $x \models \neg \phi$ iff for all $y$ with $R(x, y)$, we have $y \not\models \phi$
Definition

We change the definition of $x \models \phi$ as follows:

- $x \models \top$
- $x \not\models \bot$
- $x \models p$ iff $p \in L(x)$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
- $x \models \phi \lor \psi$ iff $x \models \phi$ or $x \models \psi$

as usual, but now:

- $x \models \neg \phi$ iff for all $y$ with $R(x, y)$, we have $y \not\models \phi$
- $x \models \phi \rightarrow \psi$ iff for all $y$ with $R(x, y)$, we have $y \models \psi$ whenever we have $y \models \phi$. 

Satisfaction in Intuitionistic Logic
Example

Let $W = \{x, y\}$, $R = \{(x, x), (x, y), (y, y)\}$, $L(x)(p) = F$, $L(y)(p) = T$. 
Example

Let $W = \{x, y\}$, $R = \{(x, x), (x, y), (y, y)\}$, $L(x)(p) = F$, $L(y)(p) = T$. Does $x \models p \lor \neg p$ hold?
Example

Let $W = \{x, y\}$, $R = \{(x, x), (x, y), (y, y)\}$, $L(x)(p) = F$, $L(y)(p) = T$. Does $x \vDash p \lor \neg p$ hold?

Since

\[ x \vDash \phi \lor \psi \iff x \vDash \phi \text{ or } x \vDash \psi \]

we would need: $x \vDash \neg p$. 
Example

Let $W = \{x, y\}$, $R = \{(x, x), (x, y), (y, y)\}$, $L(x)(p) = F$, $L(y)(p) = T$. Does $x \Vdash p \lor \neg p$ hold? Since

- $x \Vdash \phi \lor \psi$ iff $x \Vdash \phi$ or $x \Vdash \psi$

we would need: $x \Vdash \neg p$.

Since

- $x \Vdash \neg \phi$ iff for all $y$ with $R(x, y)$, we have $y \nvdash \phi$

we cannot establish $x \Vdash \neg p$. 
Example

Let \( W = \{x, y\}, \ R = \{(x, x), (x, y), (y, y)\}, \)
\( L(x)(p) = F, \ L(y)(p) = T. \) Does \( x \vdash p \lor \neg p \) hold?

Since

- \( x \vdash \phi \lor \psi \) iff \( x \vdash \phi \) or \( x \vdash \psi \)

we would need: \( x \vdash \neg p. \)

Since

- \( x \vdash \neg \phi \) iff for all \( y \) with \( R(x, y) \), we have \( y \not\vdash \phi \)

we cannot establish \( x \vdash \neg p. \)

Idea

Do not allow “assumptions”, even if they exhaust all possibilities.
Review of Modal Logic

Some Modal Logics

Natural Deduction in Modal Logic

- More Boxes
- Rules
- Extra Rules
- Example

Knowledge in Multi-Agent Systems
In addition to proof boxes for assumptions, we introduce blue boxes that express knowledge about an arbitrary accessible world.
Idea

In addition to proof boxes for assumptions, we introduce *blue boxes* that express knowledge about an *arbitrary accessible world*.

Rules about blue boxes
Idea

In addition to proof boxes for assumptions, we introduce blue boxes that express knowledge about an arbitrary accessible world.

Rules about blue boxes

- Whenever $\Box \phi$ occurs in a proof, $\phi$ may be put into a subsequent blue box.
Dashed Boxes

Idea

In addition to proof boxes for assumptions, we introduce blue boxes that express knowledge about an arbitrary accessible world.

Rules about blue boxes

- Whenever $\Box \phi$ occurs in a proof, $\phi$ may be put into a subsequent blue box.
- Whenever $\phi$ occurs at the end of a blue box, $\Box \phi$ may be put after that blue box.
Rules for □

Introduction of □:

\[ \vdash \square \phi \quad \square \phi \rightarrow [\square i] \]

\[ \vdash \square \phi \quad \square \phi \]
Elimination of □:

\[ \square \phi \]

\[ \vdash \phi \]

\[ [\square \ e] \]
Extra Rules for KT45

\[\begin{array}{c}
\Box \phi \\
\hline
\phi
\end{array} \quad [T] \quad \begin{array}{c}
\Box \phi \\
\Box \Box \phi
\end{array} \quad [4] \quad \begin{array}{c}
\neg \Box \phi \\
\Box \neg \Box \phi
\end{array} \quad [5]
\]
Example Proof

\( \vdash_K \Box p \land \Box q \rightarrow \Box (p \land q) \)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \Box p \land \Box q )</td>
<td>assumption</td>
</tr>
<tr>
<td>2</td>
<td>( \Box p )</td>
<td>( \land e_1 \ 1 )</td>
</tr>
<tr>
<td>3</td>
<td>( \Box q )</td>
<td>( \land e_2 \ 1 )</td>
</tr>
<tr>
<td>4</td>
<td>( p )</td>
<td>( \Box e_2 )</td>
</tr>
<tr>
<td>5</td>
<td>( q )</td>
<td>( \Box e_3 )</td>
</tr>
<tr>
<td>6</td>
<td>( p \land q )</td>
<td>( \land i \ 4,5 )</td>
</tr>
<tr>
<td>7</td>
<td>( \Box (p \land q) )</td>
<td>( \Box i \ 4-6 )</td>
</tr>
<tr>
<td>8</td>
<td>( \Box p \land \Box q \rightarrow \Box (p \land q) )</td>
<td>( \rightarrow i \ 1-7 )</td>
</tr>
</tbody>
</table>
Review of Modal Logic

Some Modal Logics

Natural Deduction in Modal Logic

Knowledge in Multi-Agent Systems

Motivation: The Wise Women Puzzle

Modal Logic KT45^n

Models of KT45^n

Formulation of Wise-Women Puzzle
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and two available white hats
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and two available white hats.
- Each wise woman is wise, can see other hats but not her own.
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and two available white hats
- Each wise woman is wise, can see other hats but not her own
- Queen asks first wise woman: Do you know the color of your hat.
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and \textit{two} available white hats
- Each wise woman is wise, can see other hats but not her own
- Queen asks first wise woman: Do you know the color of your hat.
  Answer: No
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and *two* available white hats
- Each wise woman is wise, can see other hats but not her own
- Queen asks first wise woman: Do you know the color of your hat.
  Answer: No
- Queen asks second wise woman: Do you know the color of your hat.
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and two available white hats.
- Each wise woman is wise, can see other hats but not her own.
- Queen asks first wise woman: Do you know the color of your hat.
  Answer: No
- Queen asks second wise woman: Do you know the color of your hat.
  Answer: No
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and two available white hats.
- Each wise woman is wise, can see other hats but not her own.
- Queen asks first wise woman: Do you know the color of your hat. Answer: No.
- Queen asks second wise woman: Do you know the color of your hat. Answer: No.
- Queen asks third wise woman: Do you know the color of your hat?
Wise Women Puzzle

- Three wise women, each wearing one hat, among three available red hats and *two* available white hats.
- Each wise woman is wise, can see other hats but not her own.
- Queen asks first wise woman: Do you know the color of your hat. Answer: No.
- Queen asks second wise woman: Do you know the color of your hat. Answer: No.
- Queen asks third wise woman: Do you know the color of your hat?
- What is her answer?
Reasoning about knowledge

We saw that KT45 can be used to reason about an agent’s knowledge.
Motivation

Reasoning about knowledge

We saw that KT45 can be used to reason about an agent’s knowledge.

Difficulty

We have three agents (queen does not count), not just one. We want them to be able to reason about each other’s knowledge.
Motivation

Reasoning about knowledge
We saw that KT45 can be used to reason about an agent’s knowledge.

Difficulty
We have three agents (queen does not count), not just one. We want them to be able to reason about each other's knowledge.

Idea
Introduce a □ operator for each agent, and a □ operator for a group of agents.
Agents

Assume a set $A = \{1, 2, \ldots, n\}$ of agents.
Agents

Assume a set $\mathcal{A} = \{1, 2, \ldots, n\}$ of agents.

Modal connectives

Replace $\square$ by:

- $K_i$ for each agent $i$
- $E_G$ for any subset $G$ of $\mathcal{A}$
## Modal Logic KT45^n

### Agents

Assume a set \( A = \{1, 2, \ldots, n\} \) of agents.

### Modal connectives

- Replace \( \Box \) by:
  - \( K_i \) for each agent \( i \)
  - \( E_G \) for any subset \( G \) of \( A \)

### Example

\[ K_1 p \land K_1 \neg K_2 K_1 p \]

means:

Agent 1 knows \( p \), and also that Agent 2 does not know that Agent 1 knows \( p \).
"Everyone knows that everyone knows"

In $\text{KT45}^n$, $E_G E_G \phi$ is stronger than $E_G \phi$. 
Common Knowledge

“Everyone knows that everyone knows”
In KT45^n, E_G E_G \phi is stronger than E_G \phi.

“Everyone knows everyone knows everyone knows”
In KT45^n, E_G E_G E_G \phi is stronger than E_G E_G \phi.
Common Knowledge

“Everyone knows that everyone knows”
In KT45^n, \(E_G E_G \phi\) is stronger than \(E_G \phi\).

“Everyone knows everyone knows everyone knows”
In KT45^n, \(E_G E_G E_G \phi\) is stronger than \(E_G E_G \phi\).

Common knowledge
The infinite conjunction \(E_G \phi \land E_G E_G \phi \land \ldots\) is called “common knowledge of \(\phi\)”, denoted, \(C_G \phi\).
Distributed Knowledge

Combine knowledge

If intelligent agents communicate with each other and use the knowledge each have, they can discover new knowledge.
Distributed Knowledge

Combine knowledge
If intelligent agents communicate with each other and use the knowledge each have, they can discover new knowledge.

Distributed knowledge
The operator $D_G \phi$ is called “distributed knowledge of $\phi$”, denoted, $D_G \phi$. 
Models of KT45\(^n\)

Definition

A model \( \mathcal{M} = (W, (R_i)_{i \in A}, L) \) of the multi-modal logic KT45\(^n\) is specified by three things:
Models of KT45^n

Definition

A model $\mathcal{M} = (W, (R_i)_{i \in \mathcal{A}}, L)$ of the multi-modal logic KT45^n is specified by three things:

1. A set $W$, whose elements are called *worlds*;
Models of KT45\(^n\)

**Definition**

A model \( \mathcal{M} = (W, (R_i)_{i \in \mathcal{A}}, L) \) of the multi-modal logic KT45\(^n\) is specified by three things:

1. A set \( W \), whose elements are called *worlds*;
2. For each \( i \in \mathcal{A} \) a relation \( R_i \) on \( W \), meaning \( R_i \subseteq W \times W \), called the accessibility relations;
Models of $KT45^n$

Definition

A model $\mathcal{M} = (W, (R_i)_{i \in \mathcal{A}}, L)$ of the multi-modal logic $KT45^n$ is specified by three things:

1. A set $W$, whose elements are called \textit{worlds};
2. For each $i \in \mathcal{A}$ a relation $R_i$ on $W$, meaning $R_i \subseteq W \times W$, called the accessibility relations;
3. A labeling function $L : W \rightarrow \mathcal{P}(\text{Atoms})$. 
Semantics of $\text{KT45}^n$

**Definition**

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:
Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $x \models p$ iff $p \in L(x)$
Semantics of $\text{KT45}^n$

**Definition**

Take a model $\mathcal{M} = (W, (R_i)_{i \in \mathcal{A}}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $x \models p$ iff $p \in L(x)$
- $x \models \neg \phi$ iff $x \not\models \phi$
Semantics of $\text{KT}45^n$

**Definition**

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $x \models p$ iff $p \in L(x)$
- $x \models \neg \phi$ iff $x \not\models \phi$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \vDash \phi$ via structural induction:

- $x \vDash p$ iff $p \in L(x)$
- $x \vDash \neg\phi$ iff $x \not\vDash \phi$
- $x \vDash \phi \land \psi$ iff $x \vDash \phi$ and $x \vDash \psi$
- $x \vDash \phi \lor \psi$ iff $x \vDash \phi$ or $x \vDash \psi$
Semantics of KT45^n

**Definition**

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $x \models p$ iff $p \in L(x)$
- $x \models \neg \phi$ iff $x \not\models \phi$
- $x \models \phi \land \psi$ iff $x \models \phi$ and $x \models \psi$
- $x \models \phi \lor \psi$ iff $x \models \phi$ or $x \models \psi$
- $x \models \phi \rightarrow \psi$ iff $x \models \psi$, whenever $x \models \phi$
- ...
Semantics of $KT45^n$ (continued)

**Definition**

Take a model $\mathcal{M} = (W, (R_i)_{i \in \mathcal{A}}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- ...
Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \Vdash \phi$ via structural induction:

- ...
- $x \Vdash K_i \phi$ iff for each $y \in W$ with $R_i(x, y)$, we have $y \Vdash \phi$
Semantics of KT45^n (continued)

Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $\ldots$
- $x \models K_i\phi$ iff for each $y \in W$ with $R_i(x, y)$, we have $y \models \phi$
- $x \models E_G\phi$ iff for each $i \in G$, $x \models K_i\phi$. 
Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $\ldots$
- $x \models K_i \phi$ iff for each $y \in W$ with $R_i(x, y)$, we have $y \models \phi$
- $x \models E_G \phi$ iff for each $i \in G$, $x \models K_i \phi$.
- $x \models C_G \phi$ iff for each $k \geq 1$, we have $x \models E_G^k \phi$. 
Semantics of $\text{KT}45^n$ (continued)

Definition

Take a model $\mathcal{M} = (W, (R_i)_{i \in A}, L)$ and a world $x \in W$. We define $x \models \phi$ via structural induction:

- $\ldots$
- $x \models K_i\phi$ iff for each $y \in W$ with $R_i(x, y)$, we have $y \models \phi$
- $x \models E_G\phi$ iff for each $i \in G$, $x \models K_i\phi$.
- $x \models C_G\phi$ iff for each $k \geq 1$, we have $x \models E^k_G\phi$.
- $x \models D_G\phi$ iff for each $y \in W$, we have $y \models \phi$, whenever $R_i(x, y)$ for all $i \in G$. 
Formulation of Wise-Women Puzzle

Setup

- Wise woman $i$ has red hat: $p_i$
- Wise woman $i$ knows that wise woman $j$ has a red hat: $K_i p_j$
Formulation of Wise-Women Puzzle

Initial situation

\[ \Gamma = \{ C(p_1 \lor p_2 \lor p_3), \]
\[ C(p_1 \rightarrow K_2 p_1), C(\neg p_1 \rightarrow K_2 \neg p_1), \]
\[ C(p_1 \rightarrow K_3 p_1), C(\neg p_1 \rightarrow K_3 \neg p_1), \]
\[ C(p_2 \rightarrow K_1 p_2), C(\neg p_2 \rightarrow K_1 \neg p_2), \]
\[ C(p_2 \rightarrow K_3 p_2), C(\neg p_2 \rightarrow K_3 \neg p_2), \]
\[ C(p_3 \rightarrow K_1 p_3), C(\neg p_2 \rightarrow K_1 \neg p_3), \]
\[ C(p_3 \rightarrow K_2 p_3), C(\neg p_2 \rightarrow K_2 \neg p_3) \} \]
Announcements

First wise woman says “No”

\[ C(\neg K_1 p_1 \land \neg K_1 \neg p_1) \]

Second wise woman says “No”

\[ C(\neg K_2 p_2 \land \neg K_2 \neg p_2) \]
First Attempt

\[ \Gamma, C(\neg K_1 p_1 \land \neg K_1 \neg p_1), C(\neg K_2 p_2 \land \neg K_2 \neg p_2) \vdash K_3 p_3 \]
First Attempt

\[ \Gamma, C(\neg K_1 p_1 \land \neg K_1 \neg p_1), C(\neg K_2 p_2 \land \neg K_2 \neg p_2) \vdash K_3 p_3 \]

Problem

This does not take time into account. The second announcement can take the first announcement into account.
Solution

Prove separately:

**Entailment 1:**

\[
\Gamma, \quad C(\neg K_1 p_1 \land \neg K_1 \neg p_1) \vdash C(p_2 \lor p_3)
\]
Solution

Prove separately:

**Entailment 1**:\[ \Gamma, C(\neg K_1 p_1 \land \neg K_1 \neg p_1) \vdash C(p_2 \lor p_3) \]

**Entailment 2**:\[ \Gamma, C(p_2 \lor p_3), C(\neg K_2 p_2 \land \neg K_2 \neg p_2) \vdash K_3 p_3 \]
Solution

Prove separately:

**Entailment 1**: 
\[ \Gamma, C(\neg K_1 p_1 \land \neg K_1 \neg p_1) \vdash C(p_2 \lor p_3) \]

**Entailment 2**: 
\[ \Gamma, C(p_2 \lor p_3), C(\neg K_2 p_2 \land \neg K_2 \neg p_2) \vdash K_3 p_3 \]

**Proof**

Through natural deduction in KT45^n.
Motivation: The Wise Women Puzzle

Modal Logic KT45

Models of KT45

Formulation of Wise-Women Puzzle

Next Week

- Hoare Logic