Modal logic

What is Modal Logic?

Narrowly, traditionally: modal logic studies reasoning that involves the use of the expressions *necessarily* and *possibly*.
More widely: *modal logic* covers a family of logics with similar rules and a variety of different symbols.

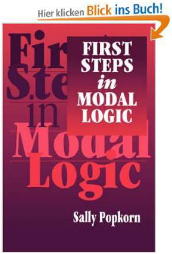
Logic	Symbols	Expressions Symbolized
Modal Logic	$\Box$	It is necessary that ...
	$\Diamond$	It is possible that ...
Deontic Logic	O	It is obligatory that ...
	P	It is permitted that ...
	F	It is forbidden that ...
Temporal Logic	G	It will always be the case that ...
	F	It will be the case that ...
	H	It has always been the case that ...
	P	It was the case that ...
Doxastic Logic	Bx	x believes that ...

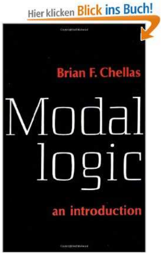
History

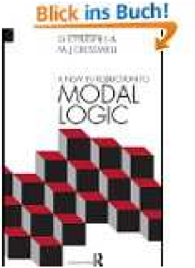
- Aristotles (356-323 BCE): developed a modal syllogistic in book I of his *Prior Analytics*, which *Theophrastus* attempted to improve
- Avicenna (980-1037): developed earliest formal system of modal logic
- William of Ockham (1287-1347) and John Duns Scotus (1266-1308): informal modal reasoning (about essence)
- C.I. Lewis (1883-1964): founded modern modal logic
- Ruth C. Barcan (1921-2012): first axiomatic systems of quantified modal logic
- Saul Kripke: Kripke semantics for modal logics; possible worldsßemantics
- A.N. Prior: created modern temporal logic in 1957
- Vaughan Pratt: introduced dynamic logic in 1976.

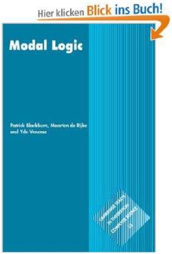
Further Reading

Garson, James, *Modal Logic*, The Stanford Encyclopedia of Philosophy (Spring 2013 Edition), Edward N. Zalta (ed.), `<http://plato.stanford.edu/archives/spr2013/entries/logic-modal/>`.









Modal Logic: Applications

Modal logics have been used in artificial intelligence applications to model

- Knowledge (including common knowledge)
- Belief (including common knowledge)
- Actions, goals, and intentions
- Ability and Obligations
- Time
- ...

There are many further applications, also in other disciplines, including philosophy, linguistics, mathematics, computer science, ... arts, poetry

(Here are some nice slides for further reading; see also the slides of Andreas Herzig)

Modal Logic: Motivation

Material implication seems actually quite unintuitive:

$$\varphi \Rightarrow \psi \text{ iff } \neg\varphi \vee \psi$$

Problem with material implication in many applications: see e.g. Dorothy Edgington's Proof of the Existence of God:

- If God does not exist, then it's not the case that if I pray, my prayers will be answered.
  $$\neg G \Rightarrow \neg(P \Rightarrow A)$$
- I do not pray: $\neg P$
- It follows: God exists. G

- If God does not exist, then it's not the case that if I pray, my prayers will be answered. $\neg g \Rightarrow \neg(p \Rightarrow a)$
- I do not pray: $\neg p$
- It follows: God exists. g

In TPTP syntax:

```
fof(ax1,axiom,((~ g) => ~ (p => a))).
fof(ax2,axiom,(~ p)).
fof(c,conjecture,(g)).
```

$$\sim g \Rightarrow \sim \underbrace{(p \Rightarrow a)}_{T}$$

T since we have $\sim p$, i.e., p is F

p	a	$p \Rightarrow a$
F	T	T
F	F	T

Hence

$$\sim g \Rightarrow \sim T$$

Hence

$$\sim\sim T \Rightarrow \sim\sim g$$

Hence

$$T \Rightarrow g$$

Hence

$$g$$

q.e.d.

modus tollens: $\varphi \Rightarrow \psi$ iff $\sim\psi \Rightarrow \sim\varphi$

why holds?

φ	ψ	$\varphi \Rightarrow \psi$	$\neg\varphi$	$\neg\psi$	$\neg\psi \Rightarrow \neg\varphi$
T	T	T	F	F	T
T	F	F	F	T	F
F	T	T	T	F	T
F	F	T	T	T	T

Modal Logic: Motivation

Lewis instead proposed the use of *strict implication*:

$$\varphi \Rightarrow \psi \text{ iff } \neg\Diamond(\varphi \wedge \neg\psi)$$

φ implies ψ iff it is not *possible* that φ and $\neg\psi$ are true.

Freie Universität Berlin

Modal Logic: Syntax

- any basic propositional symbol $p \in P$ is a modal logic formula
- if φ and ψ are modal logic formulas, then so are $\neg\varphi$, $\varphi \vee \psi$, $\varphi \wedge \psi$, and $\varphi \Rightarrow \psi$
- if φ is a modal logic formula, then so are $\Box\varphi$ and $\Diamond\varphi$

Prominent modal logics are constructed from a weak logic called K (after Saul Kripke).

Theorems of Basic Modal Logic K

- if φ is a theorem of propositional logic, then φ is also a theorem of K
- Necessitation Rule: If φ is a theorem of K, then so is $\Box\varphi$
- Distribution Axiom: $\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi)$

Modal Logics beyond K

From base logic K we can derive at other modal logics by adding further axioms

Name	*Axioms*
K	$\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi)$
$M(orT)$	$\Box\varphi \Rightarrow \varphi$
D	$\Box\varphi \Rightarrow \Diamond\varphi$
B	$\varphi \Rightarrow \Box\Diamond\varphi$
4	$\Box\varphi \Rightarrow \Box\Box\varphi$
5	$\Diamond\varphi \Rightarrow \Box\Diamond\varphi$

A variety of logics may be developed using K as a foundation by adding combinations of the above axioms.

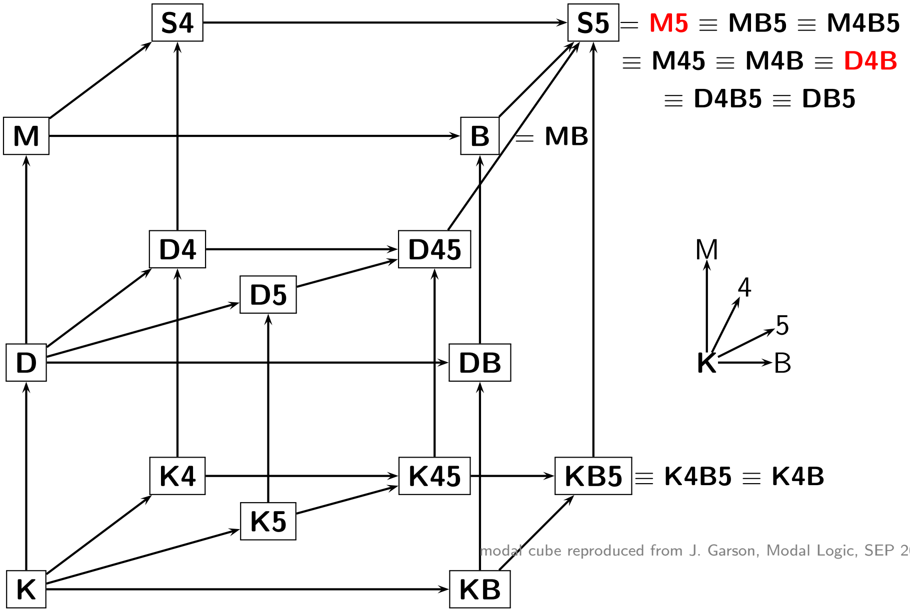
Modal Logics beyond K

Many philosophers consider logic S5 (K+M+4+5) an adequate choice for *necessity*.

In S5, $** \ldots \Box = \Box$ and $** \ldots \Diamond = \Diamond$, where each * is either $\Box$ or $\Diamond$. This amounts to the idea that strings containing both boxes and diamonds are equivalent to the last operator in the sequence. Saying that it is possible that A is necessary is the same as saying that A is necessary.

Modal logic can be extended to multi-modal logic, where the $\Box$ and $\Diamond$ operators are annotated with the identifier of the agent who has that knowledge; see wise men puzzle above.

Modal Logic Cube



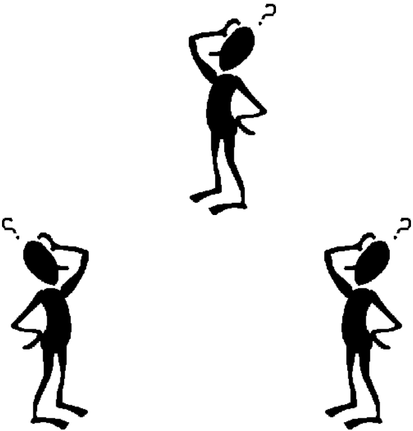
modal cube reproduced from J. Garson, Modal Logic, SEP 2009

Can you represent and solve the following problem?

Wise Men Puzzle

Once upon a time, a king wanted to find the wisest out of his three wisest men. He arranged them in a circle and told them that he would put a white or a black spot on their foreheads and that one of the three spots would certainly be white. The three wise men could see and hear each other but, of course, they could not see their faces reflected anywhere. The king, then, asked to each of them to find out the color of his own spot. After a while, the wisest correctly answered that his spot was white.

How could he know that?



Can you represent and solve the following problem?

Wise Men Puzzle

Once upon a time, a king wanted to find the wisest out of his three wisest men. He arranged them in a circle and told them that he would put a white or a black spot on their foreheads and that one of the three spots would certainly be white. The three wise men could see and hear each other but, of course, they could not see their faces reflected anywhere. The king, then, asked to each of them to find out the color of his own spot. After a while, the wisest correctly answered that his spot was white.

How could he know that?

	A	B	C	Answer
1	*	b	b	A w
2	b	*	b	B w
3	b	b	*	C w
4	*	w	b	A w (→ 2.)
5	*	b	w	A w (→ 3.)
6	w	*	b	B w (→ 1.)
7	b	*	w	B w (→ 3.)
8	w	b	*	C w (→ 1.)
9	b	w	*	C w (→ 2.)
10	*	w	w	A w (→ 3.)
11	w	*	w	B w (→ 8.)
12	w	w	*	C w (→ 4.)

Can you represent and solve the following problem?

Wise Men Puzzle

Once upon a time, a king wanted to find the wisest out of his three wisest men. He arranged them in a circle and told them that he would put a white or a black spot on their foreheads and that one of the three spots would certainly be white. The three wise men could see and hear each other but, of course, they could not see their faces reflected anywhere. The king, then, asked to each of them to find out the color of his own spot. After a while, the wisest correctly answered that his spot was white.

How could he know that?

$$\Box_{fool}\, ws\_a \vee ws\_b \vee ws\_c$$

$$\Box_{fool}\, \varphi \Rightarrow \Box_a\, \varphi$$

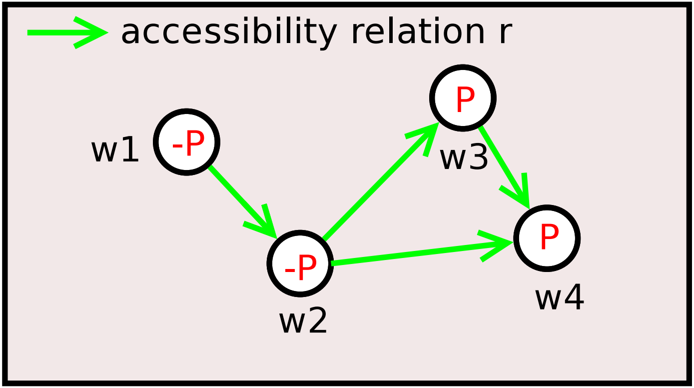
$$\Box_{fool}\, \varphi \Rightarrow \Box_b\, \varphi$$

$\ldots$

$$\neg \Box_a\, ws\_a$$

$$\neg \Box_b\, ws\_b$$

Query: $\Box_c\, ws\_c$


accessibility relation r
w1
-P
-P
w2
P
w3
P
w4

Kripke Style Semantics

A Modal Frame $F = \langle W, R, v \rangle$

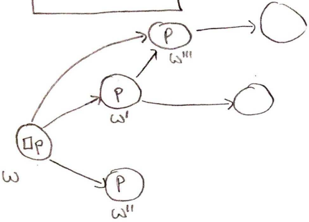
... consists of set of possible worlds W, a binary accessibility relation R between worlds, and an evaluation function v for assigning truth values to the basic propositional symbols ($v : PropSym \times W \longrightarrow \{T, F\}$).

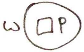
Truth of a modal formula φ for a frame F and a world w

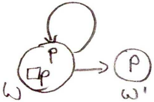
$$
\begin{array}{rll}
F, w \models p & \textit{iff} & v(p, w) \\
F, w \models \neg\varphi & \textit{iff} & F, w \not\models \varphi \\
F, w \models \varphi \vee \psi & \textit{iff} & F, w \models \varphi \text{ or } F, w \models \psi \\
F, w \models \varphi \wedge \psi & \textit{iff} & F, w \models \varphi \text{ and } F, w \models \psi \\
F, w \models \varphi \Rightarrow \psi & \textit{iff} & F, w \not\models \varphi \text{ or } F, w \models \psi \\
F, w \models \Box\varphi & \textit{iff} & F, w' \models \varphi \text{ for all } w' \text{ with } wRw' \\
F, w \models \Diamond\varphi & \textit{iff} & \text{there exists } w' \text{ with } wRw' \text{ s.t. } F, w' \models \varphi
\end{array}
$$

Kripke Style Semantics

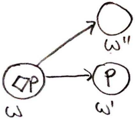
$\boxed{\mathcal{F}, \omega \models \Box p}$



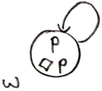




$\boxed{\mathcal{F}, \omega \models \Diamond p}$



<u>nicht möglich:</u>



Truth of a modal formula (in base modal logic K)

A modal formula φ is true (or valid) iff it is true for all frames F and all worlds w.

Exercises:

- Show that the Distribution axiom $\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi)$ is valid in logic K.
- Show that axiom T $\Box\varphi \Rightarrow \varphi$ is valid iff the accessibility relation R is reflexive.

Kripke Style Semantics

We have the following correspondences

Name	*Axioms*	Condition on R
K	$\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi)$	*none*
$M(or T)$	$\Box\varphi \Rightarrow \varphi$	*reflexive*
D	$\Box\varphi \Rightarrow \Diamond\varphi$	*serial*
B	$\varphi \Rightarrow \Box\Diamond\varphi$	*symmetric*
4	$\Box\varphi \Rightarrow \Box\Box\varphi$	*transitive*
5	$\Diamond\varphi \Rightarrow \Box\Diamond\varphi$	*euclidean*

Resolution Method

(Note: $F[A]$ means that F has a subformula A at a certain position)

Given: Formula F Goal: Show that F is valid.

Step 1: $\neg F$ (negate F, show contradiction)

Step 2: Apply clause normalization rules exhaustively

$$\frac{F[\neg\neg A]}{F[A]} \quad \frac{F[\neg(A \vee B)]}{F[\neg A \wedge \neg B]} \quad \frac{F[\neg(A \wedge B)]}{F[\neg A \vee \neg B]} \quad \frac{F[A \Rightarrow B]}{F[\neg A \vee B]} \quad \frac{F[A \Leftrightarrow B]}{F[A \Rightarrow B \wedge B \Rightarrow A]}$$

Step 3: Build resulting clause set

$$(L_1^1 \vee \ldots \vee L_1^n) \wedge \ldots \wedge (L_k^1 \vee \ldots \vee L_k^m) \rightsquigarrow \{\underbrace{L_1^1 \vee \ldots \vee L_1^n}_{C_1}, \ldots, \underbrace{L_k^1 \vee \ldots \vee L_k^m}_{C_2}\}$$

Literals: constant symbols or negated constant symbols

called clauses

Step 4: apply resolution + factorization exhaustively until empty clause is found. (empty disjunction) $\square$

$$\frac{A \vee C \quad \neg A \vee D}{C \vee D}\ \text{res} \qquad \frac{A \vee A \vee C}{A \vee C}\ \text{factorization}$$

(Note: We implicitly assume commutativity and associativity of '$\vee$' i.e.: $C \vee A \vee B \vee A \triangleq A \vee A \vee B \vee C$)

If $\square$ is found, then $\neg F$ is not satisfiable (i.e. contradictory), and hence F is valid.

If $\square$ is not found and if no further new clauses can be derived, then $\neg F$ is satisfiable and F is not valid.

Theorem Proving in Propositional Logic

Resolution Example $\quad F = ((A \Rightarrow B) \vee (A \wedge \neg(B \vee C))) \vee (A \wedge C)$

Step 1: $\neg(((A \Rightarrow B) \vee (A \wedge \neg(B \vee C))) \vee (A \wedge C))$

Step 2: $\neg((A \Rightarrow B) \vee (A \wedge \neg(B \vee C))) \wedge \neg(A \wedge C)$

$\neg(A \Rightarrow B) \wedge \neg(A \wedge \neg(B \vee C)) \wedge \neg(A \wedge C)$

$A \wedge \neg B \wedge \neg(A \wedge \neg(B \vee C)) \wedge \neg(A \wedge C)$

$A \wedge \neg B \wedge (\neg A \vee \neg\neg(B \vee C)) \wedge \neg(A \wedge C)$

$A \wedge \neg B \wedge (\neg A \vee B \vee C) \wedge \neg(A \wedge C)$

$A \wedge \neg B \wedge (\neg A \vee B \vee C) \wedge (\neg A \vee \neg C)$

Step 3: $\{A, \neg B, \neg A \vee B \vee C, \neg A \vee \neg C\}$

Step 4:

$$\dfrac{\dfrac{A \qquad \dfrac{\dfrac{A \quad \neg A \vee \neg C}{\neg C} \qquad C \vee \neg A \vee B}{\neg A \vee B}}{B} \qquad \neg B}{\square}$$

1:	A	
2:	$\neg B$	
3:	$\neg A \vee B \vee C$	
4:	$\neg A \vee \neg C$	
5:	$\neg C$	res (1+4)
6:	$\neg A \vee B$	res (5+3)
7:	B	res (1+6)
8:	$\square$	res (2+7)

Alternative Representation

Refutation for $\neg F$ found $\rightsquigarrow$ F is valid!

Freie Universität Berlin

Tableau Method

Given: Formula F **Goal:** Show that F is valid

Step 1: $\neg F$ (negate F, show contradiction)

Step 2: Apply the following rules ~~exhaustively~~ until closed paths are detected, or until no new information is obtained on paths

$$\frac{\neg\neg A}{A}\,\neg\neg \qquad \frac{A \wedge B}{\begin{array}{c}A\\B\end{array}}\,\wedge \qquad \frac{A \vee B}{A \mid B}\,\vee \qquad \frac{\neg(A \wedge B)}{\neg A \mid \neg B}\,\neg\wedge \qquad \frac{\neg(A \vee B)}{\begin{array}{c}\neg A\\ \neg B\end{array}}\,\neg\vee \qquad \frac{A \Rightarrow B}{\neg A \mid B}\,\Rightarrow \qquad \frac{\neg(A \Rightarrow B)}{\begin{array}{c}A\\ \neg B\end{array}}\,\neg\Rightarrow$$

$$\frac{A \Leftrightarrow B}{\begin{array}{c}A \Rightarrow B\\ B \Rightarrow A\end{array}}\,\Leftrightarrow \qquad \frac{\neg(A \Leftrightarrow B)}{\neg(A \Rightarrow B) \mid \neg(B \Rightarrow A)}\,\neg\Leftrightarrow$$

Step 3: Check whether Tableau is closed, that is, whether each path through the tableaux contains a complementary pair of formulas P and $\neg P$.

If closed, then $\neg F$ is not satisfiable (i.e. contradictory), and hence F is valid.

If open and if no new information can be obtained on paths, then $\neg F$ is satisfiable and F is not valid (but countersatisfiable).

Theorem Proving in Propositional Logic

Freie Universität Berlin

Tableau Example $F = ((A \Rightarrow B) \vee (A \wedge \neg(B \vee C))) \vee (A \wedge C)$

Step 1: 1: $\neg(((A \Rightarrow B) \vee (A \wedge \neg(B \vee C))) \vee (A \wedge C))$

Step 2:

2:	$\neg((A \Rightarrow B) \vee (A \wedge \neg(B \vee C)))$	$\neg\vee(1)$
3:	$\neg(A \wedge C)$	$\neg\vee(1)$
4:	$\neg(A \Rightarrow B)$	$\neg\vee(2)$
5:	$\neg(A \wedge \neg(B \vee C))$	$\neg\vee(2)$
6:	A	$\neg\Rightarrow(4)$
7:	$\neg B$	$\neg\Rightarrow(4)$

8: $\neg A$ $\neg\wedge(3)$ — $\lightning(6+8)$

9: $\neg C$ $\neg\wedge(3)$

10: $\neg A$ $\neg\wedge(5)$ — $\lightning(6+10)$

11: $\neg\neg(B \vee C)$ $\neg\wedge(5)$

12: $B \vee C$ $\neg\neg(11)$

13: B $\vee(12)$ — $\lightning(7+13)$

14: C $\vee(12)$ — $\lightning(9+14)$

Step 3 $\leadsto$ Formula $\neg F$ is contradictory, hence F is valid.

Tableau Method for Modal Logic K

Given: Modal logic formula F. **Goal:** Show that F is valid

Step 1: $\neg F$ (negate F, show contradiction) and start with $(1)\ \neg F$ ← world "prefix"

Step 2: Apply the following rules until closed paths are detected, or until no new information is obtained.

$$\frac{(\sigma)\ \neg\neg A}{(\sigma)\ A}\neg\neg \quad \frac{(\sigma)\ A \land B}{\begin{matrix}(\sigma)\ A\\ B\end{matrix}}\land \quad \frac{(\sigma)\ \neg(A \lor B)}{\begin{matrix}(\sigma)\ \neg A\\ (\sigma)\ \neg B\end{matrix}}\neg\lor \quad \frac{(\sigma)\ \neg(A \Rightarrow B)}{\begin{matrix}(\sigma)\ A\\ (\sigma)\ \neg B\end{matrix}}\neg\Rightarrow \quad \frac{(\sigma)\ A \Leftrightarrow B}{\begin{matrix}(\sigma)\ A \Rightarrow B\\ (\sigma)\ B \Rightarrow A\end{matrix}}\Leftrightarrow$$

$$\frac{(\sigma)\ \neg(A \land B)}{(\sigma)\ \neg A \mid (\sigma)\ \neg B}\neg\land \quad \frac{(\sigma)\ A \lor B}{(\sigma)\ A \mid (\sigma)\ B}\lor \quad \frac{(\sigma)\ A \Rightarrow B}{(\sigma)\ \neg A \mid (\sigma)\ B}\Rightarrow \quad \frac{(\sigma)\ \neg(A \Leftrightarrow B)}{(\sigma)\ \neg(A \Rightarrow B) \mid (\sigma)\ \neg(B \Rightarrow A)}\neg\Leftrightarrow$$

simple extension with world "prefix" (σ)

new rules for modal operators

$$\frac{(\sigma)\ \Diamond A}{(\sigma.n)\ A}\Diamond \text{ (if } (\sigma.n) \text{ is new to the branch)} \qquad \frac{(\sigma)\ \neg\Box A}{(\sigma.n)\ \neg A}\neg\Box \text{ (if } (\sigma.n) \text{ is new to the branch)}$$

$$\frac{(\sigma)\ \Box A}{(\sigma.n)\ A}\Box \text{ (if } (\sigma.n) \text{ already occurs on branch)} \qquad \frac{(\sigma)\ \neg\Diamond A}{(\sigma.n)\ \neg A}\neg\Diamond \text{ (if } (\sigma.n) \text{ already occurs on branch)}$$

Step 3: Check whether tableau is closed ... as before ... But now complementarity of formulas also requires identical prefixes: σF and $\sigma \neg F$ are complementary.

Freie Universität Berlin

$\boxed{\text{Modal Tableaux Example}}$ $F = \Box(A \land B) \Rightarrow (\Box A \land \Box B)$

in Logic K

Step 1: 1: (1) $\neg(\Box(A \land B) \Rightarrow (\Box A \land \Box B))$

Step 2: 2: (1) $\Box(A \land B)$ $\neg\Rightarrow(1)$

3: (1) $\neg(\Box A \land \Box B)$ $\neg\Rightarrow(1)$

4: (1) $\neg\Box A$ $\neg\lor(3)$	5: (1) $\neg\Box B$ $\neg\lor(3)$
6: (1.1) $\neg A$ $\neg\Box(4)$	10: (1.1) $\neg B$ $\neg\Box(5)$
7: (1.1) $A \land B$ $\Box(2)$	11: (1.1) $A \land B$ $\Box(2)$
8: (1.1) A $\land(7)$	12: (1.1) A $\land(11)$
9: (1.1) B $\land(7)$	13: (1.1) B $\land(11)$
↯ (6+8)	↯ (10+13)

Step 3: all paths are closed

For Modal Logics M, D, 4, 5, S4, S5 etc. add (combinations of) the following rules:

T	$\dfrac{(\sigma)\ \Box A}{(\sigma)\ A}\,T$	$\dfrac{(\sigma)\ \neg\Diamond A}{(\sigma)\ \neg A}\,T$
D	$\dfrac{(\sigma)\ \Box A}{(\sigma)\ \Diamond A}\,D$	$\dfrac{(\sigma)\ \neg\Diamond A}{(\sigma)\ \neg\Box A}\,D$
B	$\dfrac{(\sigma.n)\ \Box A}{(\sigma)\ A}\,B$	$\dfrac{(\sigma.n)\ \neg\Diamond A}{(\sigma.n)\ \neg A}\,B$
4	$\dfrac{(\sigma)\ \Box A}{(\sigma.n)\ \Box A}\,4$	$\dfrac{(\sigma)\ \neg\Diamond A}{(\sigma.n)\ \neg\Diamond A}\,4$
4r	$\dfrac{(\sigma.n)\ \Box A}{(\sigma)\ \Box A}\,4r$	$\dfrac{(\sigma.n)\ \neg\Diamond A}{(\sigma)\ \neg\Diamond A}\,4r$

all prefixes are required to occur already on the branch

Theorem Proving in Propositional Modal Logic

Freie Universität Berlin

How to extend the rules for further Modal Logics

Logic	added rules (added to K rules)
D	D
T	T
K4	4
B	B, 4
S4	T, 4
S5	T, 4, 4r

Example:

For theorem proving in logic S4 we thus need all rules for Logic K plus the rules T and 4

Theorem Proving in Propositional Modal Logic

Freie Universität Berlin

Modal Tableau Example in logic S4

$F = (\Box\Diamond(\Box A \Rightarrow \Box\Diamond B)) \Rightarrow (\Box A \Rightarrow \Box\Diamond B)$

Step 1: $1{:}\ (1)\ \neg((\Box\Diamond(\Box A \Rightarrow \Box\Diamond B)) \Rightarrow (\Box A \Rightarrow \Box\Diamond B))$

Step 2:

$2{:}\ (1)\ \Box\Diamond(\Box A \Rightarrow \Box\Diamond B) \quad \neg\Rightarrow(1)$

$3{:}\ (1)\ \neg(\Box A \Rightarrow \Box\Diamond B) \quad \neg\Rightarrow(1)$

$4{:}\ (1)\ \Box A \quad \neg\Rightarrow(3)$

$5{:}\ (1)\ \neg\Box\Diamond B \quad \neg\Rightarrow(3)$

$6{:}\ (1.1)\ \neg\Diamond B \quad \neg\Box(5)$

$7{:}\ (1.1)\ \Box A \quad 4(4)$

$8{:}\ (1.1)\ \Diamond(\Box A \Rightarrow \Box\Diamond B) \quad T(2)$

$9{:}\ (1.1.1)\ \Box A \Rightarrow \Box\Diamond B \quad \Diamond(8)$

$10{:}\ (1.1.1)\ \neg\Box A \quad \Rightarrow(9)$	$11{:}\ (1.1.1)\ \Box\Diamond B \quad \Rightarrow(9)$
$12{:}\ (1.1.1)\ \Box A \quad 4(7)$	$13{:}\ (1.1.1)\ \Diamond B \quad T(11)$
$\lightning\ (10+12)$	$14{:}\ (1.1.1)\ \neg\Diamond B \quad 4(6)$
	$\lightning\ (13+14)$

Step 3 all closed

Theorem Proving in Propositional Modal Logic

Exercisco:

(1) Prove $\Diamond(P \Rightarrow \Box P)$ in T

(2) Prove $(\Box P \wedge \Box Q) \Rightarrow \Box(\Box P \wedge \Box Q)$ in $K4$

(3) Prove $\Diamond \Box P \Rightarrow \Box P$ in $S5$

(4) Prove $\Box P \vee \Box \neg \Box P$ in $S5$

Theorem Proving in Propositional Modal Logic

Example:
Wise Men Puzzle

(1) The king announce that at least one man has a white spot.
→ It is common knowledge that at least one man has a white spot.

$\Box_{fool}(ws\_a \vee ws\_b \vee ws\_c)$

"every fool knows that ..."

(2) If one man has a white spot, then the other two men can see this (and hence know this).

$\Box_{fool}(ws\_a \Rightarrow \Box_b\, ws\_a)$ $\quad$ $\Box_{fool}(ws\_a \Rightarrow \Box_c\, ws\_a)$

$\Box_{fool}(ws\_b \Rightarrow \Box_a\, ws\_b)$ $\quad$ $\Box_{fool}(ws\_b \Rightarrow \Box_c\, ws\_b)$

$\Box_{fool}(ws\_c \Rightarrow \Box_a\, ws\_c)$ $\quad$ $\Box_{fool}(ws\_c \Rightarrow \Box_b\, ws\_c)$

(3) If one man does not have a white spot, then the other two men can see this (and hence know this).

$\Box_{fool}(\neg ws\_a \Rightarrow \Box_b \neg ws\_a)$ $\quad$ $\Box_{fool}(\neg ws\_a \Rightarrow \Box_c \neg ws\_a)$

$\Box_{fool}(\neg ws\_b \Rightarrow \Box_a \neg ws\_b)$ $\quad$ $\Box_{fool}(\neg ws\_b \Rightarrow \Box_c \neg ws\_b)$

$\Box_{fool}(\neg ws\_c \Rightarrow \Box_a \neg ws\_c)$ $\quad$ $\Box_{fool}(\neg ws\_c \Rightarrow \Box_b \neg ws\_c)$

(4) $\Box_{fool}$ is a modeled as an S4 operator:

$(\Box_{fool}\, \varrho) \Rightarrow \varrho$ "only things that hold can be known"

$(\Box_{fool}\, \varrho) \Rightarrow (\Box_{fool}\Box_{fool}\, \varrho)$ "if ϱ is known, then it is known that ϱ is known"

axioms not needed if we use the S4 proof rules

(5) $\Box_a, \Box_b, \Box_c$ are modeled simply a K operators (we actually do not need more to solve the problem; it would not be wrong to also model them as S4 operators)

(6) We need "Inclusion axioms" that connect common knowledge with the knowledge of the agents a, b, c.

$(\Box_{fool}\, \varrho) \Rightarrow (\Box_a\, \varrho)$ $\quad (\Box_{fool}\, \varrho) \Rightarrow (\Box_b\, \varrho)$ $\quad (\Box_{fool}\, \varrho) \Rightarrow (\Box_c\, \varrho)$

"if ϱ is common knowledge then ϱ is known by a"

(7) Whenever a wise man does (not) know something, then the others know that he does (not) know this.

$(\Box_a \varrho) \Rightarrow (\Box_b \Box_a \varrho)$ $\quad (\Box_a \varrho) \Rightarrow (\Box_c \Box_a \varrho)$

$(\Box_b \varrho) \Rightarrow (\Box_a \Box_b \varrho)$ $\quad (\Box_b \varrho) \Rightarrow (\Box_c \Box_b \varrho)$

$(\Box_c \varrho) \Rightarrow (\Box_a \Box_c \varrho)$ $\quad (\Box_c \varrho) \Rightarrow (\Box_b \Box_c \varrho)$

$(\neg\Box_a \varrho) \Rightarrow (\Box_b \neg\Box_a \varrho)$ $\quad (\neg\Box_a \varrho) \Rightarrow (\Box_c \neg\Box_a \varrho)$

$(\neg\Box_b \varrho) \Rightarrow (\Box_a \neg\Box_b \varrho)$ $\quad (\neg\Box_b \varrho) \Rightarrow (\Box_c \neg\Box_b \varrho)$

$(\neg\Box_c \varrho) \Rightarrow (\Box_a \neg\Box_c \varrho)$ $\quad (\neg\Box_c \varrho) \Rightarrow (\Box_b \neg\Box_c \varrho)$

(8) The first two wise men do not know whether they have a white spot

$\neg\Box_a$ ws_a

$\neg\Box_b$ ws_b

(9) We can now prove: The third wise man does know that he has a white spot.

$\Box_c$ ws_c

Example: Friends Puzzle

(1) Peter is a friend of John, so if Peter knows that John knows something then John knows that Peter knows the same thing. That is we know the following persistence axiom

$$(\Box_p \Box_J \varphi) \Rightarrow (\Box_J \Box_p \varphi)$$

knowledge of john

knowledge of peter

(2) $\Box_p$ and $\Box_J$ are S4 modal operators (S4 is often considered as adequate for modeling knowledge)

$$(\Box_p \varphi) \Rightarrow \varphi \qquad (\Box_J \varphi) \Rightarrow \varphi$$

$$(\Box_p \varphi) \Rightarrow (\Box_p \Box_p \varphi) \qquad (\Box_J \varphi) \Rightarrow (\Box_J \Box_J \varphi)$$

Axioms not needed if we use S4 proof rules

(3) Peter is married, so if Peter's wife knows something, then Peter knows the same thing. That is we have the following inclusion axiom:

$$(\Box_{wp} \varphi) \Rightarrow (\Box_p \varphi)$$

(4) $\Box_{wp}$ is also an S4 modality: $(\Box_{wp} \varphi) \Rightarrow \varphi \qquad (\Box_{wp} \varphi) \Rightarrow (\Box_{wp} \Box_{wp} \varphi)$

(5) John and Peter have an appointment. Let us consider the following situation:

$\Box_P$ time — "Peter knows the time of their appointment"

$\Box_P \Box_J$ place — "Peter knows that John knows the time of their appointment"

$\Box_{wp}(\Box_P \text{time} \Rightarrow \Box_J \text{time})$ — "Peter's wife knows that if Peter knows the time of their appointment, then John knows that too." (Since they are friends.)

$\Box_P(\Box_J(\text{place} \wedge \text{time})) \Rightarrow (\Box_J \text{appointment}))$ — "Peter knows that if John knows the time and place of their appointment, then John knows that they have an appointment."

(6) From these assumptions one can prove:

"Each of the friends knows that the other one knows that he has an appointment"

$(\Box_J \Box_P \text{appointment}) \wedge (\Box_P \Box_J \text{appointment})$

Exercise

Once upon a time, a king wanted to find the wisest out of his two wisest men. He told them that he would put a white or a black spot on their foreheads and that one of the two spots would certainly be white. The two wise men could see and hear each other but, of course, they could not see their faces reflected anywhere. The king, then, asked to each of them to find out the color of his own spot. After a while, the wisest correctly answered that his spot was white.

Exercise: Encode this situation in propositional modal logic.