Modal logic

Motivation

- In natural language, we often use **modes** of truth, e.g. “possibly true ”, “necessarily true”, “known to be true”, “believed to be true”, “true in the future”.
- E.g. the sentence

  Tony Blair is prime minister.

  is true, but will be false at some point in the future.

Motivation

- Consider the sentence

  There are nine planets in the solar system.

  It is is possibly true, but not necessarily true, because there might be more planets.
- The sentence

  The square root of 9 is 3.

  is necessarily true, and true in the future. But it does not enjoy all modes of truth: it may not be believed to be true (if the believer is mistaken).

Modal logic: overview

- We shall study **modal logics**, which can express modes of truth.
- Modal logics are very useful in computer science.
- E.g. it can be used to reason about the knowledge of agents.
- It can also be used to specify the behaviour of computer programs and reactive systems (e.g. CTL).

Modal formulæ

The language of basic modal logic is that of propositional logic with two extra connectives $\Box$ and $\Diamond$ ("box" and "diamond").

Definition. The formulæ of basic modal logic are defined by the following grammar:

$$A, B ::= \bot | p | A \wedge B | A \vee B | A \rightarrow B | \Box A | \Diamond A,$$

where p ranges over atomic formula.

$\Box$ and $\Diamond$

- In basic modal logic, $\Box$ and $\Diamond$ are read "box" and "diamond".
- But when we express a mode of truth, we may read them appropriately.
- E.g. in the logic for necessity and possibility, $\Box$ is read "necessarily" and $\Diamond$ "possibly".
- In the logic of agent Q's knowledge, $\Box$ is read "agent Q knows" and $\Diamond$ is read "it is consistent with agent Q's knowledge that".

Towards a semantics

- A situation for propositional logic is simply assigns a truth value to each atomic formula.
- This is not enough to compute the truth values of formulæ of the form $\Box A$ or $\Diamond A$.
- This problem is solved **Kripke models**, which were introduced by the philosopher-logician Saul Kripke.

Kripke models

Definition. A **(Kripke) model** of basic modal logic consists of

1. a set W, whose elements are called **worlds**;
2. a relation R on W (i.e. $R \subseteq W \times W$), called the **accessibility relation**;
3. a function $L : W \rightarrow P(Atoms)$, called the **labelling function**.

The labelling function describes the propositional situation in each world.

Example of a Kripke model

$$W = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

$$R = \{(x_1, x_2), (x_1, x_3), (x_2, x_2),$$
$$(x_2, x_3), (x_3, x_2), (x_4, x_5), (x_5, x_4), (x_5, x_6)\}$$

$$L(x_1) = \{q\}, L(x_2) = \{p, q\}, L(x_3) = \{p\},$$
$$L(x_4) = \{q\}, L(x_5) = \{\}, L(x_6) = \{p\}$$

Warning about terminology

Unfortunately, the meaning of “model” in the Kripke sense clashes with the definition

> “A model of a formula A is a situation that satisfies A”

that we have seen earlier.

Situations and forcing

- A situation in our sense is a pair (M, x) consisting of a Kripke model M and a world x in M.
- One usually writes $x \Vdash A$ instead of $(M, x) \models A$.
- The terminology for $x \Vdash A$ is “x **forces** A”.

Forcing for the propositional part

The forcing relation for propositional connectives looks like the satisfaction relation of classical propositional logic, except that the labelling function is needed to determine whether $x \Vdash p$:

$$\begin{array}{rll}
x \Vdash A \wedge B & \text{iff} & x \Vdash A \text{ and } x \Vdash B \\
x \Vdash A \vee B & \text{iff} & x \Vdash A \text{ or } x \Vdash B \\
x \Vdash A \rightarrow B & \text{iff} & x \Vdash A \text{ implies } x \Vdash B \\
x \nVdash \bot & & \\
x \Vdash p & \text{iff} & p \in L(x)
\end{array}$$

Forcing for $\Box$ and $\Diamond$

$x \Vdash \Box A$	iff for each $y \in W$ with $R(x, y)$ we have $y \Vdash A$
$x \Vdash \Diamond A$	iff there is a $y \in W$ with $R(x, y)$ such that $y \Vdash A$

Semantic entailment

Semantic entailment for basic modal logic is defined in the same way as for propositional logic or predicate logic, except that the situations are now of the form (M, x):

Definition. A set Γ of formulæ **semantically entails** a formula B if every situation that satisfies every formula in Γ also satisfies B. We write

$$\Gamma \models B.$$

Semantic entailment in terms of forcing

Using forcing, semantic entailment can be re-stated as follows:

Definition. A set Γ of formulæ of basic modal logic **semantically entails** a formula B if for every world x in every model M, we have $x \Vdash B$ whenever $x \Vdash A$ for every $A \in \Gamma$.

This is how the definition is normally presented.

Validity

Definition. A formula is called **valid** if it is satisfied by every situation, i.e. if

$$\models A.$$

Validity and semantic entailment

Evidently, we have

$$A_1, \ldots, A_n \models B$$

if and only if

$$\models (A_1 \wedge \ldots \wedge A_n) \to B.$$

So, studying semantic entailment is essentially the same as studying validity.

Examples

Which of the following formulæ are valid?

$$(\Box(A \to B) \wedge \Box A) \to \Box B \qquad (K)$$

$$\Box A \to A \qquad (T)$$

$$\Box A \to \Box\Box A \qquad (4)$$

$$\Diamond A \to \Box\Diamond A \qquad (5)$$

$$\Box A \to \Diamond A \qquad (D)$$

$$\Box A \vee \Box\neg A \qquad (X)$$

Embedding propositional logic

Proposition. Let Γ be a set of propositional formulæ, and let A be a propositional formula. Then

$\Gamma \models A$ in the sense of propositional logic
iff $\Gamma \models A$ in the sense of modal logic.

In other words, basic modal logic is a **conservative extension** of propositional logic.

The proposition holds essentially because the forcing semantics of the propositional connectives agrees with the standard semantics of propositional logic. The details of the proof are left to the keen.

Semantic equivalence

Definition. Formulæ A and B are called **semantically equivalent** if $A \models B$ and $B \models A$. We write

$$A \equiv B.$$

Examples

Which of the following statements are true?

$$\Diamond\neg A \equiv \neg\Box A \qquad \Box\neg A \equiv \neg\Diamond A$$

$$\Box(A \wedge B) \equiv \Box A \wedge \Box B \qquad \Diamond(A \vee B) \equiv \Diamond A \vee \Diamond B$$

$$\Box(A \vee B) \equiv \Box A \vee \Box B \qquad \Diamond(A \wedge B) \equiv \Diamond A \wedge \Diamond B.$$

Intended meanings of $\Box$

The intended meaning of $\Box A$ can be e.g.

- It is necessarily true that A.
- It will always be true that A.
- It ought to be true that A.
- Agent Q believes that A.
- Agent Q knows that A.

Getting $\Diamond$ from $\Box$

We have

$$\Diamond A \equiv \neg\Box\neg A;$$

this suggests that we obtain the intended meaning of $\Diamond$ from the intended meaning of $\Box$.

Getting $\Diamond$ from $\Box$

Quiz:

1. If $\Box A$ is “it is necessarily true that A”, then what is $\Diamond$?
2. If $\Box$ is “A will always be true”, then what is $\Diamond$?
3. If $\Box A$ is “it ought to be that A”, then what is $\Diamond$?
4. If $\Box A$ is “agent Q believes A”, then what is $\Diamond$?
5. If $\Box A$ is “agent Q knows A”, then what is $\Diamond$?

Which formulæ should be valid?

	T	4	5	D	K	X
A is necessarily true	y	y	y	y	y	n
A will always be true	?	y	n	n	y	n
It ought to be that A	n	?	n	y	y	n
Agent Q believes that A	n	y	y	y	y	?
Agent Q knows that A	y	?	y	y	?	?

(T) $\Box A \rightarrow A$

(4) $\Box A \rightarrow \Box\Box A$

(5) $\Diamond A \rightarrow \Box\Diamond A$

(D) $\Box A \rightarrow \Diamond A$

(K) $(\Box(A \rightarrow B) \wedge \Box A) \rightarrow \Box B$

(X) $\Box A \vee \Box\neg A$

Which formulæ should be valid?

In many cases, the answers are debatable! For example, we must clarify

- whether the present is part of the future,
- whether believers have an opinion on every matter,
- whether, in the case of "knowledge", we assume **positive introspection** (4), **negative introspection**, and **logical omniscience** (K).

Meaning of the accessibility relation

$\Box A$	$R(x, y)$
A will always be true	y is in the future of x
It ought to be that A	y is acceptable according to the information at x
A is necessarily true	y is possible according to the information at x
Agent Q knows that A	y could be the actual world according to Q's knowledge at x
Agent Q believes that A	y could be the actual world according to Q's beliefs at x

Conditions for R

To make sense w.r.t. a particular meaning (future, knowledge, etc.), the accessibility relation R may have to satisfy extra conditions. E.g. R is called

- **reflexive** if, for every $x \in W$, we have $R(x, x)$;
- **transitive** if, for every $x, y, z \in W$, it holds that $R(x, y)$ and $R(y, z)$ imply $R(x, z)$;
- **serial** if, for every $x \in W$, there is a y such that $R(x, y)$.
- **Euclidean** if, for every $x, y, z \in W$ with $R(x, y)$ and $R(x, z)$, we have $R(y, z)$.

Example: knowledge

Recall that xRy means "y could be the actual world according to Q's knowledge at x".

- Quiz: should R be reflexive?
- If we assume positive introspection, then R should be transitive.

  To see this, let xRy and yRz. We have xRz if z could be the actual world according to Q's knowledge at x. This is the case if any fact A known at x is true at z—formally, if $x \Vdash \Box A$ implies $z \Vdash A$ for all A. To see this, $x \Vdash \Box A$. By positive introspection, we have $x \Vdash \Box\Box A$. Because xRy, we have $y \Vdash \Box A$. Because yRz, we have $z \Vdash A$.

Correspondence theory

There is a close correspondence between axioms like $T, 4, 5$, and D and the aforementioned conditions for R:

- R is reflexive iff every Kripke model based on R satisfies T.
- Same for "transitive" and 4.
- Same for "Euclidean" and 5.
- Same for "serial" and D.

Proof. $\Rightarrow$: lecture/exercise; $\Leftarrow$: see Huth/Ryan if interested.