Modal Logics

Prof. Dr. Jürgen Dix
Dr. Nils Bulling

Time: Tuesday 10–12, Thursday: 10–12
Place: Am Regenbogen, SR 210

Website

http://www.in.tu-clausthal.de/abteilungen/cig/cigroot/teaching

Visit regularly!

There you will find important information about the lecture, documents, labs et cetera.

Lecture: Prof. J. Dix, Dr. N. Bulling
Exam: tba

Outline

Acknowledgement

This lecture was first given in 2008 by Jürgen Dix, Nils Bulling, and Michael Köster. The current version is based on that but has been extended.

Parts of this course (in particular parts of Chapters 2 and 5) were inspired by the course given by Carlos Areces and Patrick Blackburn at ESSLLI 2008.

We would like to thank Hans van Ditmarsch, who provided us with his slides from a lecture at ESSLLI 2008.

What is this lecture about? (1)

We give an overview on modal logics, which can be used in many areas of computer science. Emphasis is on theoretical-logical aspects, not on applications.

After a recap of propositional and predicate logic in Chapter 1, we introduce modal languages and the main technical notion, Kripke frames in Chapter 2. The important idea is to relate properties of the **accessibility relation** in frames to **axioms** in the modal language. We also discuss the relationship between modal logic and predicate logic (basic hybrid logic and the standard translation). In Chapter 3, we introduce the basic modal logics (normal modal logics) as axiomatic calculi, we state sound- and completeness results (using canonical models), and discuss the finite model property for modal systems.

What is this lecture about? (2)

In Chapter 4, we apply our theories to epistemic logics (to reason about **knowledge**), and introduce public announcement logic (where **common knowledge** can be described). We also investigate MSPASS, a modal logic theorem prover.

We consider in more advanced topics. While we have seen before that many properties can be expressed as modal formulae, we show in Chapter 5 that even more cannot be expressed. To this end, we consider $w_1 \leftrightsquigarrow w_2$, the notion of modally equivalent worlds, and its relation to bisimulation, which leads to the Hennessy-Milner theorem and van Benthem's characterization of standard translation of basic modal formulae.

Blackburn, P., de Rijke, M. and Venema, Y. (2001).
Modal Logic
Cambridge Tracts in Theoretical Computer Science 53
Cambridge University Press.

Boolos, G. (1979)
The Unprovability of Consistency.
Cambridge University Press.

van Ditmarsch, H., van der Hoek, W. and Kooi, B. (2008)
Dynamic Epistemic Logic.
Springer.

1. Introduction

Content of this Chapter

We present a modern view to modal logic. After a short overview about important historical events, we define what a logic is. Then, we introduce the semantical perspective on logic and show that there is more than one logic. After demonstrating the use of relational structures we specify our first logic, propositional logic, and describe a calculus to do inferences. We end with an introduction to predicate logic (also called first-order logic).

1.1 Overview

The Journey to Modal logic

The first to invent a modal logic was Aristotle.

Figure: Aristotle
(384 BC - 322 BC)

He developed the syllogisms and the idea of **necessity** and **possibility**:

Example 1.1

p is possible if $\neg p$ is not necessary.

Birth of Modal Logics

In 1918 Clarence Irving Lewis published his Survey of Symbolic Logic.

Figure: C.I. Lewis (1883 - 1964)

Improvements by Kurt Gödel

However, Kurt Gödel first proposed the standard system for classical propositional logic with the rule of generalization. He also proved the completeness of logic, solving a famous problem of Hilbert.

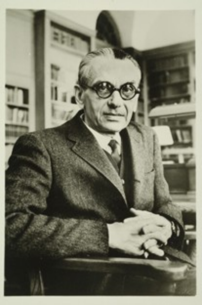


Figure: Kurt Gödel (1906 - 1978)

What is a Logic?

We present a framework for thinking about logics as:

- languages for describing a problem,
- ways of talking about relational structures and models.

These are the two key components in the way we will approach logic:

1. Language:
   fairly simple, precisely defined, formal languages.
2. Model: or relational structure
   simple “world” that the logic talks about.

Semantical Perspective

Our perspective on logic is fundamentally *semantical*, i.e. the *meaning* of an expression.

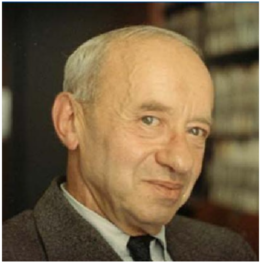


Figure: Alfred Tarski (1902-1983)

This perspective is often called model-theoretic perspective or Tarskian perspective.

Modal Logic Fundamentals

Saul Aaron Kripke's papers *A Completeness Theorem in Modal Logic* and *Semantical Considerations on Modal Logic* was a breakthrough in the 60ies: **“Londres est jolie.”**

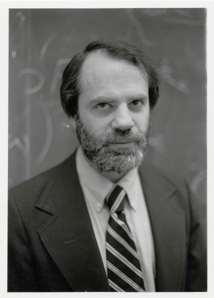


Figure: Saul Aaron Kripke (1940 -)

Logic or Logics

The semantical perspective leads us to the following question:
How many logics are there?

1. *The Monotheistic Approach*
   Choose one of all possible logical languages as the only one.
2. *The Polytheistic Approach*
   Investigate different logical languages.

Monotheism in the 20th Century

Logical monotheism was a powerful force for much of the twentieth century.

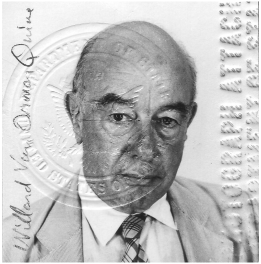


Figure: Willard van Orman Quine (1908-2000)

First-order classical logic is the one and only logic we need.

Polytheism in the 21st Century

- Polytheism gradually became the dominant thread as time went by.
- Logic found its way into:
  - computer science,
  - artificial intelligence,
  - economics and
  - cognitive science as well as
  - natural language semantics.

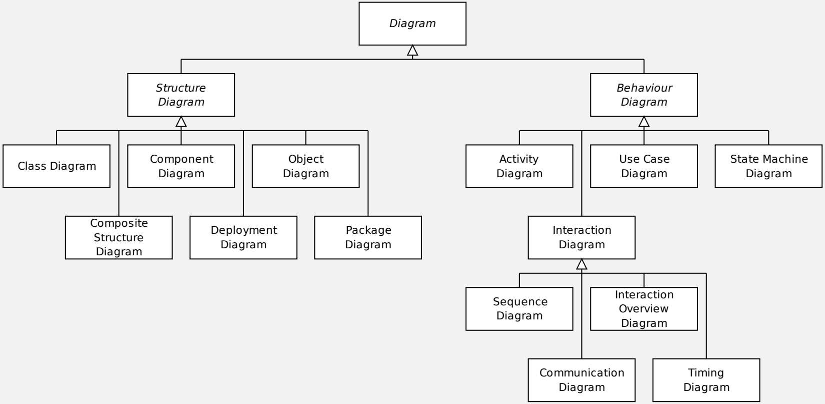
Polytheism and Semantics

The *semantical perspective* gives us an natural way to look at polytheism:

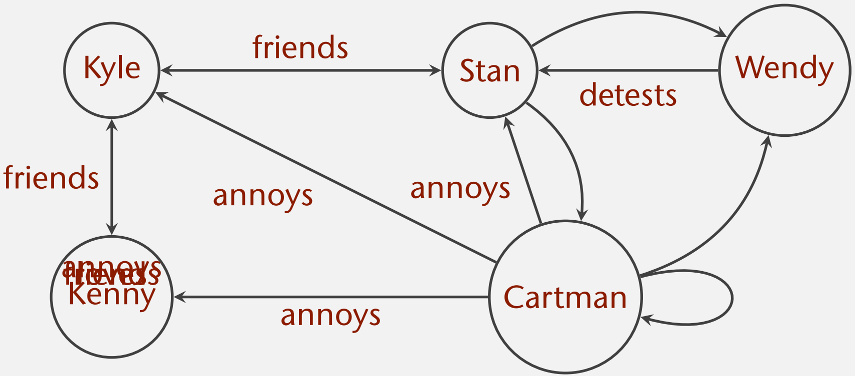
- Once we have fixed the model or relational structures (our "world") it becomes natural to play with different ways of talking about it.
- Indeed one can talk about relational structures without bothering too much about the logic at all.

What are Relational Structures?

Example 1.2 (UML Class Diagram)

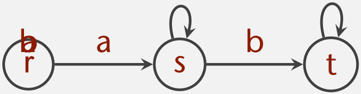


Example 1.3 (South Park Relational Structure)



Example 1.4 (Finite State Automaton for $a^n b^m$)

Given the formal language $a^n b^m$ with $n, m > 0$. The relational structure is then:



Here, r is the start state and t is the only final state.

Three Big Topics

In this course we use the **model-theoretic perspective** to provide a window on the following issues:

Inference: The methods for deriving new information by working with the logic.

Expressivity: What can be described using this logic?

Computation: Which price do we have to pay for this expressivity?

Muddy Children

Example 1.5 (Muddy Children)

- A group of playing children is called back by their father. They gather around him.
- Some of them have become dirty:
  1. they may have mud on their forehead,
  2. children can only see whether others are muddy,
  3. and not if there is any mud on their own forehead.
- All this is commonly known, and the children are perfect logicians.
- Father: “At least one of you has mud on his or her forehead.”
- Father: “Will those who know whether they are muddy please step forward.”
- If nobody steps forward, father keeps repeating the request.

1.2 Propositional Logic

Syntax

The *propositional language* is built upon

Propositional symbols: $p, q, r, \ldots, p_1, p_2, p_3, \ldots$

Logical connectives: $\neg$ and $\vee$

Grouping symbols: $($, $)$

Often we consider only a finite, nonempty set of propositional symbols and refer to it as $\mathcal{P}rop$.

Logical connectives are used to construct formulae from the propositional symbols.

Definition 1.6 (Propositional Language $\mathcal{L}_{PL}$)

The **propositional language** $\mathcal{L}_{PL}(\mathcal{P}rop)$ (over the set of propositions $\mathcal{P}rop$) is given by the formulae defined by the following grammar

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi$$

where $\mathsf{p} \in \mathcal{P}rop$.

Example 1.7

Blackboard.

In the following we assume that $\mathcal{P}rop$ is fixed and omit it.

What about the other connectives? They are defined as **macros**.

Definition 1.8 (Macros)

We define the following syntactic constructs as macros (p $\in \mathcal{P}rop$):

$$
\begin{aligned}
\bot &:= \mathrm{p} \wedge \neg \mathrm{p} \\
\top &:= \neg \bot \\
\varphi \wedge \psi &:= \neg(\neg\varphi \vee \neg\psi) \\
\varphi \rightarrow \psi &:= \neg\varphi \vee \psi \\
\varphi \leftrightarrow \psi &:= (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)
\end{aligned}
$$

Example 1.9

Blackboard.

Semantics

How to determine whether a propositional logic formula is true?

Intuitively, we have: (**t** stands for *true*, **f** for *false*)

$\top$		is **t**
$\bot$		is **f**
$\neg\varphi$ is **t**	iff	φ is **f**
$\varphi \vee \psi$ is **t**	iff	φ or ψ (or both) are **t**
$\varphi \wedge \psi$ is **t**	iff	φ and ψ are **t**
$\varphi \rightarrow \psi$ is **t**	iff	either φ is **f** or ψ is **t**
$\varphi \leftrightarrow \psi$ is **t**	iff	φ and ψ are both **t**, or both **f**

Firstly, we fix the truth value of propositional symbols.

Definition 1.10 (Valuation)

A valuation (or truth assignment) $v : \mathcal{P}rop \to \{\mathsf{t}, \mathsf{f}\}$ for a language $\mathcal{L}_{PL}(\mathcal{P}rop)$ is a mapping from the set of propositional constants defined by $\mathcal{P}rop$ into the set $\{\mathbf{t}, \mathbf{f}\}$.

A valuation fixes the value of individual propositional variables.

But we are particularly interested in the truth of formulae like $(\mathsf{p} \vee \mathsf{q}) \wedge \mathsf{r}$.

So, it remains to define the semantics of the Boolean connectives.

Again, we do only give the semantics for the basic connectives. The semantics for the macros is derived from these basic cases.

Definition 1.11 (Semantics $v \models \varphi$)

Let v be a valuation. Inductively, we define the notion of a formula φ being true or satisfied by v (notation: $v \models \varphi$):

$v \models \mathrm{p}$ iff $v(\mathrm{p}) = \mathbf{t}$ and $\mathrm{p} \in \mathcal{P}rop$,

$v \models \neg\varphi$ iff not $v \models \varphi$,

$v \models \varphi \vee \psi$ iff $v \models \varphi$ or $v \models \psi$

Given a set $\Sigma \subseteq \mathcal{L}_{PL}$ we write $v \models \Sigma$ iff $v \models \varphi$ for all $\varphi \in \Sigma$. We use $v \not\models \varphi$ instead of not $v \models \varphi$.

Exercise: Write down the appropriate clauses for the other connectives previously defined as macros.

Truth Tables

Truth tables are a conceptually simple way of working with PL (invented by Wittgenstein in 1918).

p	q	$\neg p$	$p \vee q$	$p \wedge q$	$p \rightarrow q$	$p \leftrightarrow q$
t	**t**	**f**	**t**	**t**	**t**	**t**
f	**t**	**t**	**t**	**f**	**t**	**f**
t	**f**	**f**	**t**	**f**	**f**	**f**
f	**f**	**t**	**f**	**f**	**t**	**t**

Fundamental Semantical Concepts

- If it is possible to find some valuation v that makes φ true, then we say φ is satisfiable.
- If φ is true for all valuations v then we say that φ is valid and write $\models \varphi$. φ is also called tautology.
- A theory is a set of formulae: $T \subseteq \mathcal{L}_{PL}$.
- A theory T is called consistent if there is a valuation v with $v \models T$; or, analogously, if for all valuations v we have that $v \not\models T$.
- A theory T is called complete if for each formula φ in the language, either $\varphi \in T$ or $\neg\varphi \in T$.

How do these definitions relate to truth tables?

Example 1.12 (Satisfiability)

Is $p \wedge \neg b$ satisfiable?

Example 1.13 (Validity (Tautology))

$a \vee \neg a$

Definition 1.14 (Propositional Tautologies)

The *set of all tautologies of propositional logic* is denoted by $\mathcal{T}aut_{PL}$.

The latter will be important in Section 2.4 where we show how new formulae can be derived from old ones in the modal calculus.

Model

- **Remember**: The valuation fixes the truth values of the individual propositional variables. By doing this the valuation describes our "world": The state we want to model.
- Taking the truth table into account we can say, each row is a possible state or a possible world. Speaking in model-theoretic terms, each row is a possible model.

Therefore we can interpret a valuation v as a model $\mathfrak{M}$ and use them synonymously.

Models Are Complete!

An extremely important property of a model (valuation) is that it makes each formula either true or false.

Lemma 1.15 (Models Are Complete)

For any model (valuation) v and any formula ϕ over any $\mathcal{P}rop$, the following holds

$$\textit{Either } v \models \phi \quad \textit{or} \quad v \models \neg\phi$$

Proof.

Blackboard □

Semantical Consequence (1)

Given a theory T we are interested in the following question: **Which facts can be derived from T?** We can distinguish two approaches:

1. semantical consequences, and
2. syntactical inference.

In this section we are mainly concerned with the first notion.

Definition 1.16 (ϕ Follows Semantically From T)

Let T be a theory and φ be a formula. We say that φ is a *semantical consequence of* T if for all valuations v:

$$v \models T \text{ implies } v \models \varphi.$$

Semantical Consequence (2)

The analogue of Lemma 1.15 is **not true for theories T**. In fact, for most theories we have

- neither $T \models \phi$
- nor $T \models \neg\phi$.

Syntactical Consequence

If new facts can be derived from old ones in an **algorithmic** fashion, we say that such a fact is *syntactically derivable*.

What do we need to provide? How to derive new facts?

We need *inference rules* which allow to derive new facts from old ones.

Consider the follwing example:

Example 1.17

We know that Bill is drunk and that if Bill is drunk he should not drive.

Can we derive that Bill should not drive?

In the previous example we used the modus ponens rule:

$$(MP) \qquad \frac{\varphi \qquad \varphi\rightarrow\psi}{\psi}$$

where φ, ψ are arbitrarily complex formulae.

On top of the line are the preconditions; below the line is the conclusion. If the preconditions are fulfilled the conclusion can be derived.

φ and ψ are not formulae of the language, but schemata. They can be replaced by any well-formed formula of the underlying language. In Definition 3.10 we will be more precise about that.

Example 1.18

Blackboard.

Expressiveness

What can we express with this language?

The Tweety and Friends Problem

Example 1.19

You want to throw a party for Tweety, his friend Gentoo and Tux. Unfortunately, they have different circles of friends and dislike different penguins. Therefore, Tweety tells you that he would like to invite either his friend the King or to exclude Gentoo's Adelie. But Gentoo proposes to invite Adelie or Humboldt or both. Tux, however, does not like Humboldt and the King too much, so he suggests to exclude at least one of them.

- Can we model this using propositional logic?
- What do we gain by doing that?
- What kind of questions can we "ask" our model?

Propositional Symbols

K := invite The King $\qquad \neg K$:= exclude The King

A := invite Adelie $\qquad \neg A$:= exclude Adelie

H := invite Humboldt $\qquad \neg H$:= exclude Humboldt

Problem Formalized

Tweety: $K \lor \neg A$:= invite The King or exclude Adelie, but not both: $\neg(K \land \neg A)$,

Gentoo: $A \lor H$:= invite Adelie or Humboldt or both,

Tux: $\neg H \lor \neg K$:= exclude Humboldt or The King or both.

Resulting Formula

$$\varphi = (K \lor \neg A) \land \neg(K \land \neg A) \land (A \lor H) \land (\neg H \lor \neg K)$$

K	A	H	$K \vee \neg A$	$\neg(K \wedge \neg A)$	$A \vee H$	$\neg H \vee \neg K$	φ
t	t	t	t	t	t	f	f
t	t	f	t	t	t	t	t
t	f	t	f	t	t	f	f
t	f	f	t	f	f	t	f
f	t	t	f	t	t	t	f
f	t	f	f	t	t	t	f
f	f	t	t	t	t	t	t
f	f	f	t	t	f	t	f

Problem

Truth tables have 2^n rows, where n is the number of propositional symbols.

Graph Coloring Problem

Problem

Given a fixed number k of colors and a graph $G = (N, E)$, where N is the set of nodes and E the set of edges.

Question

Can we color each node so that

- we only use the k colors, and
- no two adjacent vertices have the same color?

Can we express this problem as a set of formulae?

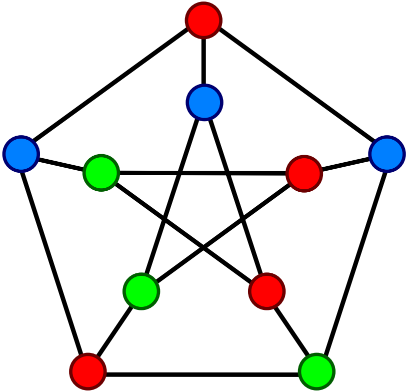


Figure: Petersen Graph

Propositional Symbols: $\mathcal{P}rop_{G,k}$

Given a graph $G = (N, E)$ and a $k \in \mathbb{N}$, we consider $|N| \times k$ propositional symbols that we write as p_{ij}.

Meaning: Node i has color j

Problem Formalized: T_G

1. Each node has one color: $p_{i1} \vee \ldots \vee p_{ik}$ for $1 \leq i \leq |N|$
2. Each node has no more than one color: $\neg p_{il} \vee \neg p_{im}$ for $1 \leq i \leq |N|$ and $1 \leq l \lneqq m \leq k$
3. Adjacent nodes have different colors: $\neg p_{il} \vee \neg p_{jl}$ for i and j adjacent and $1 \leq l \leq k$

Relating Graphs and Sets of Formulae

Lemma 1.20 (Formalizing Graphs)

For each graph $G = (N, E)$ and $k \in \mathbb{N}$, we consider the language $\mathcal{P}rop_{G,k}$ and the set of formulae $T_{G,k}$. Then the **k-colorings of $G = (N, E)$** *are in one-to-one correspondence with the* **models of $T_{G,k}$** *over the set of propositions $\mathcal{P}rop_{G,k}$.*
In particular: **there is a k-coloring** *of $G = (N, E)$ if and only if the set* **$T_{G,k}$ is satisfiable.**

1.3 Inferences

Inference Tasks: The Three Questions

Given a certain logical description φ. The semantical perspective allows us to think about the following inference tasks:

Inference Tasks

Model Checking: Given φ and a model $\mathfrak{M}$, does the formula correctly describe this model?

Satisfiability Checking: Given φ, does there exist a model in which the formula is true?

Validity Checking: Given φ, is it true in all models?

Model Checking

- This is the simplest of the three tasks.
- Nevertheless it is useful, e.g. for hardware verification.

Example 1.21

Think of a model $\mathfrak{M}$ as a **mathematical picture** of a chip. A logical description φ might define some security issues. If $\mathfrak{M}$ fulfills φ, then this means that the **chip will be secure** wrt. these issues.

Model checking for PL

Given a valuation v and a description φ. Is $v \models \varphi$ true?

Example 1.22

Given a model v in which a is **t** and b is **f**.
Is $\varphi := (\mathsf{a} \vee \neg \mathsf{b}) \rightarrow \mathsf{b}$ true?

- $(\top \vee \neg \bot) \rightarrow \bot$
- $(\top \vee \top) \rightarrow \bot$
- $\top \rightarrow \bot$
- $\bot$

Complexity

Exercise.

Satisfiability Checking

- One can interpret the description as a constraint.
- Is there anything that matches this description?
- We have to create model after model until we find one that satisfies φ.
- In the worst case we have to generate all models.

Example 1.23

Given a sudoku. Is there a solution and how does it look like?
One has to describe (maybe not actually build) all models, i.e. all possible configurations, until one is found that fulfills all constraints.

Satisfiability Checking for PL

Given a description φ. Is there a valuation v s.t. $v \models \varphi$ is true?

Complexity

How complex is that?

Validity Checking

- The intuitive idea is that we write down a set of all our fundamental and indisputable axioms.
- Then, with this theory, we derive new formulae.

Example 1.24

Mathematics: We are usually given a set of axioms.
E.g. Euclid's axioms for geometry. We want to prove whether a certain statement follows from this set, or can be derived from them.

Validity Checking for PL

Given a description φ. Does $v \models \varphi$ hold for all valuations v?

Complexity

How complex is that?

Satisfiability Checking With the Tableaux Method

- A tableau is a tree-like structure to visualize attempts to create a model.
- For building a tableau there exist some rules to systematically split the input formula into subformulae.
- Each branch of this tree represents a way of trying to build a model:
  - closed branches don't lead to models,
  - open branches do.

Rules

For splitting the input formula into subformulae we need the following rules:

$\wedge$ $\quad \dfrac{\varphi \wedge \psi}{\begin{array}{c}\varphi\\ \psi\end{array}}$ $\qquad \neg 1$ $\quad \dfrac{\varphi}{\neg\neg\varphi}$

$\vee$ $\quad \dfrac{\varphi \vee \psi}{\varphi \mid \psi}$ $\qquad \neg 2$ $\quad \dfrac{\neg\neg\varphi}{\varphi}$

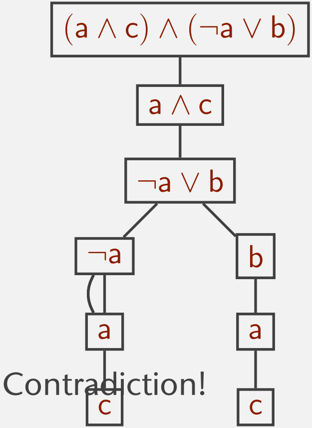
Rules for $\rightarrow, \leftrightarrow$

Exercise.

Example 1.25 (Satisfiability Checking)

Given the **input formula** $(a \wedge c) \wedge (\neg a \vee b)$. Is there a **valuation** v s.t. this formula holds?

$\wedge$	$\dfrac{\varphi \wedge \psi}{\substack{\varphi \\ \psi}}$
$\vee$	$\dfrac{\varphi \vee \psi}{\varphi \mid \psi}$
$\neg 1$	$\dfrac{\varphi}{\neg\neg\varphi}$
$\neg 2$	$\dfrac{\neg\neg\varphi}{\varphi}$



Validity Checking

Theorem 1.26 (Satisfiability, Validity are Dual)

A formula φ is valid iff $\neg\varphi$ is not satisfiable.

$\leadsto$ We can use the tableaux method for checking validity.

Proof

Blackboard.

Example 1.27

Given formula $(a \land b) \to a$. Is this formula valid?

formula: $(a \land b) \to a \equiv \neg(a \land b) \lor a$

negated: $\neg(\neg(a \land b) \lor a) \equiv \neg\neg(a \land b) \land \neg a$

$\neg 2$ and $\land$: $a \land b$ and $\neg a$

$\land$: a and b

Contradiction! $\Rightarrow \neg(a \land b) \lor a$ is valid!

SPASS

- an Automated Theorem Prover for First-Order Logic with Equality,
- from MPI, Saarbrücken

www.spass-prover.org

"If you are interested in sex, drugs, rock'n roll or fish [...], you may be disappointed by the performance of SPASS. If you are interested in first-order logic theorem proving, the formal analysis of software, [...] modal logic theorem proving, SPASS may offer you the right functionality."

DFG-Syntax

The syntax of an input file has to be of the following type:

```
problem ::= "begin_problem(" identifier ")."
            description
            logical_part
            settings*
            "end_problem."
```

The overall structure of a problem therefore contains a description about a problem, the formalisation and further settings.

The logical part is mainly defined by:

```
logical_part ::= [symbol_list]
                 [declaration_list]
                 formula_list*
                     etc...
```

While the first two are optional the `formula_list` has to describe at least one formula.

The `symbol_list` is constructed as follows:

```
symbol_list ::= "list_of_symbols."
                functions*
                predicates*
                   etc...
                "end_of_list."
```

This means, symbols can either be functions or predicates.

```
formula_list ::= "list_of_formulae(" type ")."
                 ("formula(" [term] [label] ").")*
                 "end_of_list."
type  ::= "axioms" | "conjectures"
```

Thus, a formula can be an axiom or a conjecture.

Example 1.28 (TweetyGentooTux (1))

```
begin_problem(TweetyGentooTux).
  list_of_descriptions.
    name({* TweetyGentooTux *}).
    author({* Michael Koester *}).
    status(unsatisfiable).
    description({* - Tweety and Friends - *}).
  end_of_list.
```

Example 1.29 (TweetyGentooTux (2))

```
  list_of_symbols.
    predicates[(Adelie,0), (Humboldt,0), (King,0)].
  end_of_list.
  list_of_formulae(axioms).
    formula(or(King,not(Adelie))).
    formula(or(Adelie,Humboldt)).
    formula(or(not(Humboldt),not(King))).
  end_of_list.
  list_of_formulae(conjectures).
    formula(and(Adelie,King)).
  end_of_list.
end_problem.
```

Example 1.30 (TweetyGentooTux (3))

```
--------------------------SPASS-START--------
SPASS beiseite: Completion found.

The saturated set of worked-off clauses is :
7[0:MRR:3.1,6.0] || King*+ -> .
5[0:MRR:4.0,2.1] || Adelie*+ -> .
6[0:MRR:1.1,5.0] ||  -> Humboldt*.
--------------------------SPASS-STOP--------
```

1.4 Predicate Logic

Motivation

For the formal description of a problem we often need a higher expressiveness than what the propositional language offers.

Example 1.31 (Continuity)

For the definition of the continuity we need:

- Variables,
- functions,
- relations, and
- quantifiers.

$$\varphi := \forall \epsilon > 0\ \exists \delta > 0\ (\forall x\ (\mid x - x_0 \mid < \delta) \rightarrow \mid f(x) - f(x_0) \mid < \epsilon)$$

Predicate logic

In addition to the propositional language (on which the modal language is built as well), the first-order language (FOL) contains variables, function-, and predicate symbols.

Definition 1.32 (Variable)

A variable is a symbol of the set $\mathcal{V}ar$. Typically, we denote variables by $x_0, x_1, \ldots$.

Example 1.33

$$\varphi := \exists x_0 \forall x_1 (P_0^2(f_0^1(x_0), x_1) \wedge P_2^1(f_1^0))$$

Functions

Definition 1.34 (Function Symbols)

Let $k \in \mathbb{N}_0$. The set of k-ary function symbols is denoted by $\mathcal{F}unc^k$. Elements of $\mathcal{F}unc^k$ are given by $f_1^k, f_2^k \dots$. Such a symbol takes k arguments. The set of all function symbols is defined as

$$\mathcal{F}unc := \bigcup_k \mathcal{F}unc^k$$

An 0-ary function symbol is called constant.

$$\varphi := \exists x_0 \forall x_1 (P_0^2(f_0^1(x_0), x_1) \wedge P_2^1(f_1^0))$$

Predicates

Definition 1.35 (Predicate Symbols)

Let $k \in \mathbb{N}_0$. The set of k-ary predicate symbols (or relation symbols) is given by $\mathcal{P}red^k$. Elements of $\mathcal{P}red^k$ are denoted by $P_1^k, P_2^k \ldots$. Such a symbol takes k arguments. The set of predicate symbols is defined as

$$\mathcal{P}red := \bigcup_k \mathcal{P}red^k$$

An 0-ary predicate symbol is called (atomic) proposition.

$$\varphi := \exists x_0 \forall x_1 (P_0^2(f_0^1(x_0), x_1) \wedge P_2^1(f_1^0))$$

Syntax

The first-order language with equality $\mathcal{L}_{FOL}$ is built from terms and formulae.

In the following we fix a set of variables, function-, and predicate symbols.

Definition 1.36 (Term)

A term over $\mathcal{F}unc$ and $\mathcal{V}ar$ is inductively defined as follows:

1. Each variable from $\mathcal{V}ar$ is a term.
2. If $t_1, \ldots t_k$ are terms then $f^k(t_1, \ldots, t_k)$ is a term as well, where f^k is an k-ary function symbol from $\mathcal{F}unc^k$.

$$\varphi := \exists x_0 \forall x_1 (P_0^2(f_0^1(x_0), x_1) \land P_2^1(f_1^0))$$

Definition 1.37 (Language)

The first-order language with equality $\mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{F}unc, \mathcal{P}red)$ is defined by the following grammar:

$$\varphi ::= P^k(t_1, \ldots, t_k) \mid \neg\varphi \mid \varphi \vee \varphi \mid \exists x(\varphi) \mid t \doteq r$$

where $P^k \in \mathcal{P}red^k$ is a k-ary predicate symbol and $t_1, \ldots, t_k$ and t, r are terms over $\mathcal{V}ar$ and $\mathcal{F}unc$.

Definition 1.38 (Macros)

We define the following syntactic constructs as macros ($P \in \mathcal{P}red^0$):

$$
\begin{aligned}
\bot &:= P \wedge \neg P \\
\top &:= \neg\bot \\
\varphi \wedge \psi &:= \neg(\neg\varphi \vee \neg\psi) \\
\varphi \rightarrow \psi &:= \neg\varphi \vee \psi \\
\varphi \leftrightarrow \psi &:= (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi) \\
\forall x(\varphi) &:= \neg\exists x(\neg\varphi)
\end{aligned}
$$

Notations

- We will often leave out the index k in f_i^k and P_i^k indicating the arity and just write f_i and P_i.
- Variables are also denoted by $u, v, w, \ldots$
- Function symbols are also denoted by $f, g, h, \ldots$
- Constants are also denoted by $a, b, c, \ldots, c_0, c_1, \ldots$
- Predicate symbols are also denoted by $P, Q, R, \ldots$
- We will use our standard notation p for 0-ary predicate symbols and also call them (atomic) propositions.

Example 1.39

Let d be a constant, $g \in \mathcal{F}unc^2$, and $f \in \mathcal{F}unc^3$, and $P \in \mathcal{P}red^2$.
Which of the following strings are terms?

1. $g(d, d)$
2. $P(d)$
3. $f(x, g(y, z), d)$
4. $g(x, f(y, z), d)$
5. $f(x, g(y, z), d)$
6. $g(f(g(d, x), g(f(x, a, z), y), f(x, y, g(x, y))), a)$
7. $d \wedge f(x)$
8. $\neg f(x, y)$

Binding Strength

We assume an order between the binding strength of Boolean connectives. The order is as follows:

1. $\neg, \forall x, \forall y$ bind most tightly;
2. then $\wedge$ and $\vee$;
3. then $\rightarrow$.

Example 1.40 (From Semigroups to Rings)

We consider $\mathcal{L}_{FOL}(\{x, y, z\}, \{0, 1, +, \cdot\}, \{\leq\})$, where $0, 1$ are constants, $+, \cdot$ are binary function symbols and $\leq$ a binary relation. What can one express in this language?

Ax 1:	$\forall x \forall y \forall z$	$x + (y + z) \doteq (x + y) + z$
Ax 2:	$\forall x$	$(x + 0 \doteq 0 + x) \ \wedge \ (0 + x \doteq x)$
Ax 3:	$\forall x \exists y$	$(x + y \doteq 0) \ \wedge \ (y + x \doteq 0)$
Ax 4:	$\forall x \forall y$	$x + y \doteq y + x$
Ax 5:	$\forall x \forall y \forall z$	$x \cdot (y \cdot z) \doteq (x \cdot y) \cdot z$
Ax 6:	$\forall x \forall y \forall z$	$x \cdot (y + z) \doteq x \cdot y + x \cdot z$
Ax 7:	$\forall x \forall y \forall z$	$(y + z) \cdot x \doteq y \cdot x + z \cdot x$

Axiom 1 describes a semigroup, the axioms 1-2 describe a monoid, the axioms 1-3 a group, and the axioms 1-7 a ring.

Example 1.41

Formalize the following sentence:

Every student x is younger than some teacher y.

- $S(x)$: x is a student
- $T(x)$: x is a teacher
- $Y(x, y)$: x is younger than y

$$\forall x(S(x) \rightarrow (\exists y(T(y) \land Y(x, y))))$$

Example 1.42

Formalize the following sentence:

Not all birds can fly.

- $B(x)$: x is a bird
- $F(x)$: x can fly

$$\neg(\forall x(B(x) \rightarrow F(x)))$$

Can we do it without negation?

$$\exists x(B(x) \land \neg F(x))$$

Example 1.43

Formalize the following sentence:

Andy and Paul have the same maternal grandmother.

- a: Andy
- p: Paul
- $M(x, y)$: x is y's mother

$$\forall x \forall y \forall u \forall v (M(x, y) \wedge M(y, a) \wedge M(u, v) \wedge M(v, p) \rightarrow x \doteq u)$$

Semantics

Definition 1.44 (Model, Structure)

A model or structure for FOL over $\mathcal{V}ar, \mathcal{F}unc$ and $\mathcal{P}red$ is given by $\mathfrak{M} = (U, I)$ where

1. U is a non-empty set of elements, called universe or domain and
2. I is called interpretation. It assigns to each function symbol $f^k \in \mathcal{F}unc^k$ a function $I(f^k) : U^k \to U$, to each predicate symbol $P^k \in \mathcal{P}red^k$ a relation $I(P^k) \subseteq U^k$; and to each variable $x \in \mathcal{V}ar$ an element $I(x) \in U$.

We write:

1. $\mathfrak{M}(P^k)$ for $I(P^k)$,
2. $\mathfrak{M}(f^k)$ for $I(f^k)$, and
3. $\mathfrak{M}(x)$ for $I(x)$.

Note that a structure comes with an interpretation I, which is based on functions and predicate symbols and assignments of the variables. But these are also defined in the notion of a language. Thus we assume from now on that the structures are compatible with the underlying language: The arities of the functions and predicates must correspond to the associated symbols.

Example 1.45

$\varphi := Q(x) \vee \forall z(P(x, g(z))) \vee \exists x(\forall y(P(f(x), y) \wedge Q(a)))$

- $U = \mathbb{R}$
- $I(a) : \{\emptyset\} \to \mathbb{R}$, $\emptyset \mapsto \pi$ constant functions,
- $I(f) : I(f) = \sin : \mathbb{R} \to \mathbb{R}$ and $I(g) = \cos : \mathbb{R} \to \mathbb{R}$,
- $I(P) = \{(r, s) \in \mathbb{R}^2 : r \leq s\}$ and $I(Q) = [3, \infty) \subseteq \mathbb{R}$,
- $I(x) = \frac{\pi}{2}$, $I(y) = 1$ and $I(z) = 3$.

Definition 1.46 (Value of a Term)

Let t be a term and $\mathfrak{M} = (U, I)$ be a model. We define inductively the **value of t wrt $\mathfrak{M}$**, written as $\mathfrak{M}(t)$, as follows:

$\mathfrak{M}(x) := I(x)$ for a variable $t = x$,

$\mathfrak{M}(t) := I(f^k)(\mathfrak{M}(t_1), \ldots, \mathfrak{M}(t_k))$ if $t = f^k(t_1, \ldots, t_k)$.

Definition 1.47 (Semantics)

Let $\mathfrak{M} = (U, I)$ be a model and $\varphi \in \mathcal{L}_{FOL}$. φ is said to be **true in $\mathfrak{M}$**, written as $\mathfrak{M} \models \varphi$, if the following holds:

$\mathfrak{M} \models P^k(t_1, \ldots t_k)$ iff $(\mathfrak{M}(t_1), \ldots, \mathfrak{M}(t_k)) \in \mathfrak{M}(P^k)$

$\mathfrak{M} \models \neg\varphi$ iff not $\mathfrak{M} \models \varphi$

$\mathfrak{M} \models \varphi \vee \psi$ iff $\mathfrak{M} \models \varphi$ or $\mathfrak{M} \models \psi$

$\mathfrak{M} \models \exists x(\varphi)$ iff $\mathfrak{M}^{[x/a]} \models \varphi$ for some $a \in U$ where $\mathfrak{M}^{[x/a]}$ denotes the model equal to $\mathfrak{M}$ but $\mathfrak{M}^{[x/a]}(x) = a$.

$\mathfrak{M} \models t \doteq r$ iff $\mathfrak{M}(t) = \mathfrak{M}(r)$

Given a set $\Sigma \subseteq \mathcal{L}_{FOL}$ we write $\mathfrak{M} \models \Sigma$ iff $\mathfrak{M} \models \varphi$ for all $\varphi \in \Sigma$.

Example 1.48

$\varphi := Q(x) \vee \forall z(P(x, g(z))) \vee \exists x(\forall y(P(f(x), y) \wedge Q(a)))$

- $U = \mathbb{R}$
- $I(a) : \{\emptyset\} \to \mathbb{R}$, $\emptyset \mapsto \pi$ constant functions,
- $I(f) : I(f) = \sin : \mathbb{R} \to \mathbb{R}$ and $I(g) = \cos : \mathbb{R} \to \mathbb{R}$,
- $I(P) = \{(r, s) \in \mathbb{R}^2 : r \leq s\}$ and $I(Q) = [3, \infty) \subseteq \mathbb{R}$,
- $I(x) = \frac{\pi}{2}$, $I(y) = 1$ and $I(z) = 3$.

Is φ true in $\mathfrak{M} = (U, I)$?

- $\mathfrak{M} \not\models Q(x)$ because $\mathfrak{M}(Q) = [3, \infty)$ and $\mathfrak{M}(x) = \frac{\pi}{2} \notin [3, \infty)$,
- $\mathfrak{M} \not\models \forall z(P(x, g(z)))$ because $(\mathfrak{M}(x), \mathfrak{M}(g)(\mathfrak{M}(z))) = (\frac{\pi}{2}, \cos(u)) \notin \mathfrak{M}(P) = \{(r, s) \in \mathbb{R}^2 : r \leq s\}$ and $\cos(u) < \frac{\pi}{2}$ for all u.
- $\mathfrak{M} \models Q(a)$ because $\mathfrak{M}(a) = \pi \in \mathfrak{M}(Q) = [3, \infty)$ is satisfied.
- $\mathfrak{M} \models \exists x(\forall y(P(f(x), y)))$ because not for all $u' \in \mathbb{R}$ exists a $u \in \mathbb{R}$ such that $(\mathfrak{M}(f)(\mathfrak{M}(x)), \mathfrak{M}(y)) = (\sin(u), u') \in \mathfrak{M}(P) = \{(r, s) \in \mathbb{R}^2 \mid r \leq s\}$ is fulfilled (e.g. $u' = -2$).

$\Rightarrow \mathfrak{M} \not\models \varphi$

FOL without function symbols

In the next chapter we extend ML such that it will be as expressive as FOL.

In order to give a clear presentation of the translation it is beneficial to first syntactically restrict FOL as much as possible without losing expressiveness.

We will show that it is enough to consider a FOL language without function symbols but constants.

How do we simulate function symbols?

$\rightsquigarrow$ blackboard

Let $T \subseteq \mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$ be a first-order theory (set of formulae).

We encode function symbols as predicate symbols. This seems to be straightforward as functions can be seen *as a special case* of relations.

Example 1.49 (Function $+$ as Relation $\oplus$)

We consider the **binary function** $+ : \mathbb{N}^2 \to \mathbb{N}, \langle n, n' \rangle \mapsto n + n'$. We can easily write this as a ternary relation $\oplus \subseteq \mathbb{N}^3$:

$$\oplus(n, m, l) \text{ iff } n + m = l.$$

Let $f^k \in \mathcal{F}unc^k$ be a function symbol of arity $k > 0$ occuring in some predicate symbol $P^l \in \mathcal{P}red^l$:

$$P^l(t_1, \ldots, t_{i-1}, f^k(t'_1, \ldots, t'_k), t_{i+1}, \ldots, t_l)$$

where $1 \leq i \leq l$. Now, we replace $P^l(\ldots)$ by the following conjunction:

$$\forall y (F(t'_1, \ldots, t'_k, y) \rightarrow P^l(t_1, \ldots, t_{i-1}, y, t_{i+1}, \ldots, t_l))$$

where $y \notin \mathcal{V}ar$ is a new variable and $F \notin \mathcal{P}red$ is a new predicate symbol.

Assume that there are n occurrences of f^k in T. If we apply the previous procedure to each of these n function symbols in which language can T be formulated?

This procedure allows to express the theory T over

$$\mathcal{L}_{FOL}(\mathcal{V}ar \cup \{y_1, \ldots, y_n\}, \mathcal{P}red \cup \{F\}, \mathcal{F}unc \setminus \{f^k\})$$

What did we achieve?

1. Removal of f^k
2. Tradeoff: n new variables and one new predicate symbol

We did not yet address the soundness of the translation!

By now, there is no formal relation between F and f^k!
E.g. what happens if $\mathfrak{M}(F)$ is not a function?

Firstly, we need to specify that F represents a function. So, we have to add the following "function axiom":

$$\forall x_1 \ldots \forall x_k \exists ! y F(x_1, \ldots, x_k, y)$$

Note that such an axiom is added to a given theory for each newly introduced predicate symbol F.

Remark 1.50 (Macro for "there is exactly one")

$\exists ! x Q$ denotes that there is exactly one x; i.e.

$$\exists ! x Q := \exists x \forall y (Q(x) \land (Q(y) \rightarrow x \doteq y))$$

Following this procedure recursively allows to translate the theory T over $\mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$ to a new theory T' over $\mathcal{L}_{FOL}(\mathcal{V}ar', \mathcal{P}red', \mathcal{F}unc^0)$ where

1. $\mathcal{V}ar' := \mathcal{V}ar \cup \mathcal{V}ar^1 \cup \cdots \cup \mathcal{V}ar^k$ where

$$\mathcal{V}ar^i := \bigcup_{f_j^i \in \mathcal{F}unc^i \cap T} \{y_{j,1}^i, \ldots, y_{j,n_{f_j^i}}^i\}$$

   and $n_{f_j^i}$ denotes the number of occurences of f_j^i in T.

2. $\mathcal{P}red' := \mathcal{P}red \cup \{F_i^{j+1} \mid f_i^j \in \mathcal{F}unc^j \cap T\}$ for $1 \leq j \leq k$.

We formally define the translations given of the previous slides.

Definition 1.51 (Translations tr_{vars}, tr_{preds})

Let T be a theory over $\mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$. We define the following functions:

- $tr_{\text{vars}}(T, \mathcal{V}ar) = \mathcal{V}ar \cup \mathcal{V}ar^1 \cup \cdots \cup \mathcal{V}ar^k$ where

$$\mathcal{V}ar^i := \bigcup_{f_j^i \in \mathcal{F}unc^i \cap T} \{y_{j,1}^i, \ldots, y_{j,n_{f_j^i}}^i\}$$

  and $n_{f_j^i}$ denotes the number of occurences of f_j^i in T.

- $tr_{\text{preds}}(T, \mathcal{P}red) = \mathcal{P}red \cup \{F_i^{j+1} \mid f_i^j \in \mathcal{F}unc^j \cap T\}$ for $1 \leq j \leq k$.

Definition 1.52 (Translation function $tr_{\textsf{theory}}$)

Let T be a theory over $\mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$. Then, $tr_{\textsf{theory}}(T)$ denotes the theory over

$$\mathcal{L}_{FOL}(tr_{\textsf{vars}}(T, \mathcal{V}ar), tr_{\textsf{preds}}(T, \mathcal{P}red), \mathcal{F}unc^0)$$

which is constructed by *recursively* replacing function symbols by the procedure described above.

Definition 1.53 ($Ax(T)$)

Let T be a theory over $\mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$. By $Ax(T)$ we denote the set of all "function axioms" (as described above), one for each for each (new) predicate symbol in $\mathcal{P}red \cap tr_{\textsf{preds}}(T, \mathcal{P}red)$.

Example 1.54

Consider the theory $T = \{P(f(r(x)))\}$ over $\mathcal{L}_{FOL}(\{x\}, \{P\}, \{f, r\})$.
Firstly, we remove function symbol f yielding the theory T' consisting of the following formulae:

1. $\forall y_1(F(r(x), y_1) \rightarrow P(y_1))$
2. $\forall x_1 \exists! y_1(F(x, y_1))$

Secondly, we remove the function symbol r yielding theory T'':

1. $\forall y_1(\forall y_2(R(x, y_2) \rightarrow F(y_2, y_1)) \rightarrow P(y_1))$
2. $\forall x_1 \exists! y_1(F(x, y_1))$
3. $\forall x_1 \exists! y_2(R(x, y_2))$

Now, we have that $T'' \subseteq \mathcal{L}_{FOL}(\{x, y_1, y_2\}, \{P, F, R\}, \emptyset)$.

Before we prove the main result of this section we state a lemma which is essential for its proof.

Definition 1.55 (Model Update tr_{model})

Let T be a theory, and $\mathfrak{M}$ a model over $\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc$. According to the previous construction we modify $\mathfrak{M}$ to a model $\mathfrak{M}' = tr_{\text{model}}(\mathfrak{M})$ over $tr_{\text{vars}}(T, \mathcal{V}ar)$, $tr_{\text{preds}}(T, \mathcal{P}red)$, and $\mathcal{F}unc^0$ such that $\mathfrak{M}'$ interprets each new predicate symbol F arisen from a function symbol $f \in \mathcal{F}unc^k$ as follows:

$$\mathfrak{M}'(F) = \{(u_1, \ldots, u_{k+1}) \mid u_{k+1} = I(f)(u_1, \ldots, u_k)\}$$

The interpretation of the new variables can be any element: It does not matter as each occurrence is bound.

Lemma 1.56 (Functions Can Be Eliminated)

Let $T \subseteq \mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$ *be a* theory *and* $\mathfrak{M}$ be a compatible model. *Then,* $tr_{\mathsf{theory}}(T)$ *is a theory over*

$$\mathcal{L}_{FOL}(tr_{\mathsf{vars}}(T, \mathcal{V}ar), tr_{\mathsf{preds}}(T, \mathcal{P}red), \mathcal{F}unc^0)$$

and it holds that

$$\mathfrak{M} \models T \quad \text{iff} \quad tr_{\mathsf{model}}(\mathfrak{M}) \models tr_{\mathsf{theory}}(T) \cup Ax(T).$$

Proof: ⇝ Exercise (structural induction).

Note, that the new theory does not contain any function symbols but constants.

This theorem shows that function symbols are not necessary. Everything can equivalently be encoded by predicate symbols.

Example 1.57

Again, consider the theory $T = \{P(f(r(x)))\}$ over $\mathcal{L}_{FOL}(\{x\},\{P\},\{f,r\})$ and model $\mathfrak{M} = (\mathbb{N}, I)$ where

- $\mathfrak{M}(f) : \mathbb{N} \to \mathbb{N}$, $\mathfrak{M}(f)(x) = x + 1$
- $\mathfrak{M}(r) : \mathbb{N} \to \mathbb{N}$, $\mathfrak{M}(r)(x) = 2x$
- $\mathfrak{M}(P) = \{2x + 1 \mid x \in \mathbb{N}\}$
- $\mathfrak{M}(x) = 2$

Then, the new model $\mathfrak{M}' := tr_{\mathsf{model}}(\mathfrak{M})$ is given as follows:

- $\mathfrak{M}'(F) = \{(x, x+1) \mid x \in \mathbb{N}\} \subseteq \mathbb{N}^2$,
- $\mathfrak{M}'(R) = \{(x, 2x) \mid x \in \mathbb{N}\} \subseteq \mathbb{N}^2$,
- $\mathfrak{M}'(P) = \mathfrak{M}(P)$
- $\mathfrak{M}(x) = 2$, and $\mathfrak{M}(y_1)$, $\mathfrak{M}(y_2)$ arbitrary.

It holds that $\mathfrak{M} \models T$ **iff** $\mathfrak{M}' \models T''$.

Theorem 1.58 (Function Free FOL)

Let $T \cup \{\varphi\} \subseteq \mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$ *be a theory. Then,* $tr_{\textsf{theory}}(T \cup \{\varphi\})$ *is a theory over*

$$\mathcal{L}_{FOL}(tr_{\textsf{vars}}(T, \mathcal{V}ar), tr_{\textsf{preds}}(T, \mathcal{P}red), \mathcal{F}unc^0)$$

and it holds that

$$T \models \varphi \quad \textit{iff} \quad tr_{\textsf{theory}}(T) \cup Ax(T \cup \{\varphi\}) \models tr_{\textsf{theory}}(\{\varphi\})$$

Proof sketch (1).

We have to show the following

$$\forall \mathfrak{M}(\mathfrak{M} \models T \quad \Rightarrow \mathfrak{M} \models \varphi)$$

iff

$$\forall \mathfrak{M}'(\mathfrak{M}' \models tr_{\text{theory}}(T) \cup Ax(T \cup \{\varphi\}) \quad \Rightarrow \mathfrak{M}' \models tr_{\text{theory}}(\{\varphi\}))$$

where models $\mathfrak{M}$ range over models over $\mathcal{V}ar, \mathcal{P}red$, and $\mathcal{F}unc$ and models $\mathfrak{M}'$ over $tr_{\text{vars}}(T, \mathcal{V}ar)$,$tr_{\text{preds}}(T, \mathcal{P}red)$, and $\mathcal{F}unc^0$. □

Proof sketch (2).

“$\Rightarrow$”: Assume $\mathfrak{M}' \models tr_{\text{theory}}(T) \cup Ax(T \cup \{\varphi\})$ and $\mathfrak{M}' \not\models tr_{\text{theory}}(\{\varphi\})$.

Let $\mathfrak{M}$ be a model with $tr_{\text{model}}(\mathfrak{M}) = \mathfrak{M}'$. (Such a model has to exist because of the function axioms!) According to the previous lemma, we have that $\mathfrak{M} \models T$ and $\mathfrak{M} \not\models \varphi$. Contradiction!

“$\Leftarrow$”: Assume that the right direction holds and that there is a model $\mathfrak{M}$ with $\mathfrak{M} \models T$ and $\mathfrak{M} \not\models \varphi$. According to the previous lemma we also have that
$tr_{\text{model}}(\mathfrak{M}) \models tr_{\text{theory}}(T) \cup Ax(T \cup \{\varphi\})$ and
$tr_{\text{model}}(\mathfrak{M}) \not\models tr_{\text{theory}}(\{\varphi\}))$ which contradicts the assumption. □

2. Languages and Semantics

2 Languages and Semantics

- Basic Modal Language $\mathcal{L}_{BML}$
- Modal Language $\mathcal{L}_{\mathrm{ML}}$
- Extending ML to FOL
- The Standard Translation

Content of this Chapter

Firstly, we talk about relational structures in detail and introduce the **basic modal language**. Secondly, we extend this language by multiple relations: This leads us to the modal languages. If we extend the modal logic further, we get a new logic equivalent to first order logic: Basic Hybrid Logic. This raises the question as **to which class of first-order formulae does the basic modal logic correspond**? The answer can be given with the standard translation.

2.1 Basic Modal Language $\mathcal{L}_{BML}$

What is a Logic?

We present a framework for thinking about logics as:

- languages for describing a problem,
- ways of talking about relational structures and models.

These are the two key components in the way we will approach logic:

1. Language:
   fairly simple, precisely defined, formal languages.
2. Model (or relational structure):
   simple “world” that the logic talks about.

Various modal logics

- knowledge → **epistemic logic**,
- beliefs → **doxastic logic**,
- obligations → **deontic logic**,
- actions → **dynamic logic**,
- time → **temporal logic**,
- and combinations of the above.

Most famous **multimodal logics**:
BDI logics of beliefs, desires, intentions (and time).

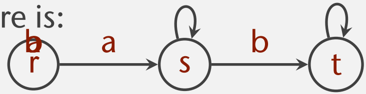
Relational Structures

A relational structure is given by $(W, \{\mathcal{R}_1, \ldots, \mathcal{R}_n\})$ and consists of:

- A non-empty set W, the elements of which are our objects of interest. They are called points, states, nodes, worlds, times, instants or situations.
- A non-empty set $\{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$ of relations, $\mathcal{R}_i \subseteq W \times W$.

Example 2.1 (Finite state automaton for $a^n b^m$)

Given the formal language $a^n b^m$ with $n, m > 0$. The relational structure is:



Here, r is the start state and t is the only final state.

The Basic Modal Language

- Standard propositional logic can be seen as a one-point relational structure.
- But relational structures can describe much more. We can talk about points, lines etc.
- Therefore, we introduce the basic modal language.

We build the basic modal language on top of the propositional language by extending $\mathcal{L}_{PL}(\mathcal{P}rop)$ with two new operators:

Possibility and necessity

$\Diamond\ \varphi$: φ is possible
(We see one or more states where φ holds.)

$\Box\ \varphi$: φ is necessary
(In all reachable states φ holds.)

A Language for Relational Structures

Definition 2.2 (Basic modal language $\mathcal{L}_{\mathrm{BML}}$)

Let $\mathcal{P}rop$ be a set of propositions. The basic modal language $\mathcal{L}_{\mathrm{BML}}(\mathcal{P}rop)$ consists of all formulae defined by the following grammar:

$$\varphi ::= \mathrm{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid \Diamond\, \varphi$$

where $\mathrm{p} \in \mathcal{P}rop$.

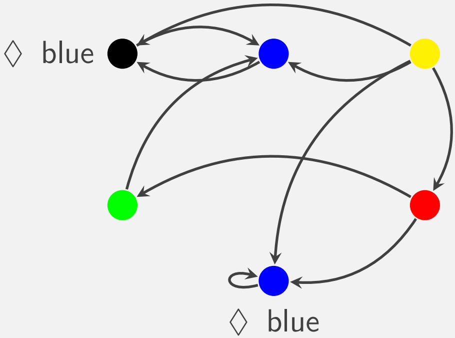
Boolean macros are defined in the standard way.
Additionally, we have the dual $\Box$ (called "box") of $\Diamond$:

$$\Box\varphi := \neg\Diamond\, \neg\varphi$$

We can talk about attributes by adding labels to nodes (e.g. painting them in a particular color).

Example 2.3 (Colored graph I)

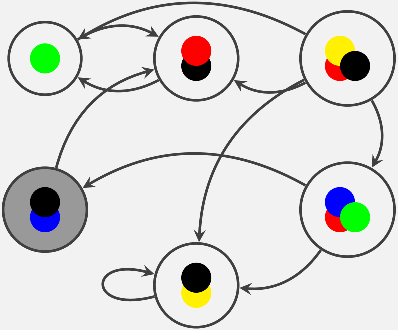
Imagine standing in a node of a colored graph. What can we see?



Example 2.4 (Colored graph II)

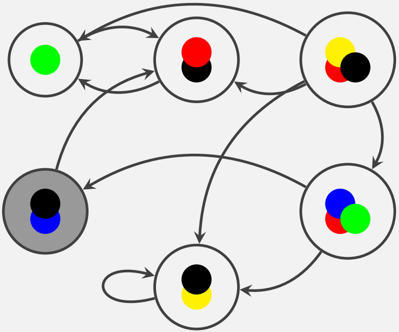
We imagine standing in a node of a colored graph. What can we see?

$\Diamond\,(\text{black} \land \text{red}) \;\land \Diamond \Diamond\, \text{green}$



Colored graph II

Example 2.5



blue $\rightarrow \Box$black

green $\rightarrow \Box$black

yellow $\rightarrow \Diamond$ yellow

Definition 2.6 (Kripke frame)

A Kripke frame is given by $\mathfrak{F} = (W, \mathcal{R})$ where

- W is a non-empty set, called set of domains or worlds,
- $\mathcal{R} \subseteq W \times W$ is a binary relation.

Frames are mainly used to talk about validities: They stand for a whole set of models.

Definition 2.7 (Kripke model)

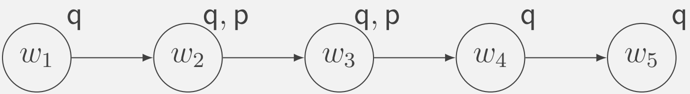
A Kripke model is given by $\mathfrak{M} = (W, \mathcal{R}, V)$ where

- $(W, \mathcal{R})$ is a Kripke frame,
- $V : \mathcal{P}rop \to \mathcal{P}(W)$ is called labelling function or valuation. We also use $V : W \to \mathcal{P}(\mathcal{P}rop)$.

Kripke frames (resp. models) are simply relational structures (resp. with labels)!

Example 2.8

Consider the frame $\mathfrak{F} = (\{w_1, w_2, w_3, w_4, w_5\}, \mathcal{R})$ where $\mathcal{R}w_iw_j$ iff $j = i + 1$ and $V(\mathsf{p}) = \{w_2, w_3\}$, $V(\mathsf{q}) = \{w_1, w_2, w_3, w_4, w_5\}$, $V(\mathsf{r}) = \emptyset$.



Frames vs. Models?

Frames

Mathematical pictures of ontologies that we find interesting. That is, frames define the fundamental structure of the domain of interest.

For example, we model time as a collection of points ordered by a strict partial order.

Models

Frames are extended by contingent information. That is, models extend the mathematical structure provided by frames by additional information.

Can Kripke models be used to interpret the propositional language?

Formal semantics of $\mathcal{L}_{ML}$.

Definition 2.9 (Semantics $\mathfrak{M}, w \models \varphi$)

Let $\mathfrak{M}$ be a Kripke model, $w \in W_{\mathfrak{M}}$, and $\varphi \in \mathcal{L}_{ML}$. φ is said to be locally true or satisfied in $\mathfrak{M}$ and world w, written as $\mathfrak{M}, w \models \varphi$, if the following holds:

$\mathfrak{M}, w \models \mathsf{p}$ iff $w \in V_{\mathfrak{M}}(\mathsf{p})$ and $\mathsf{p} \in \mathcal{P}rop$,

$\mathfrak{M}, w \models \neg\varphi$ iff not $\mathfrak{M}, w \models \varphi$

$\mathfrak{M}, w \models \varphi \vee \psi$ iff $\mathfrak{M}, w \models \varphi$ or $\mathfrak{M}, w \models \psi$

$\mathfrak{M}, w \models \Diamond\varphi$ iff there is a world $w' \in W$ such that Rww' and $\mathfrak{M}, w' \models \varphi$

Given a set $\Sigma \subseteq \mathcal{L}_{ML}$ we write $\mathfrak{M}, w \models \Sigma$ iff $\mathfrak{M}, w \models \varphi$ for all $\varphi \in \Sigma$.

What about $\Box\varphi$? $\rightsquigarrow$ blackboard

Internal and Local

Satisfaction of formulae is internal and local!

Internal: Formulae are evaluated inside models at some given world.

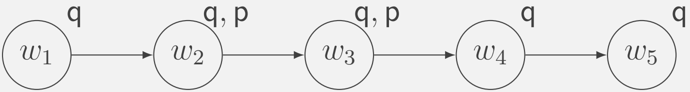
Local: Given a world it is only possible to refer to direct sucessors of this world.

How does first-order logic compare to that?

Some Examples

Example 2.10

$\mathfrak{F} = (\{w_1, w_2, w_3, w_4, w_5\}, \mathcal{R})$ where $\mathcal{R}w_iw_j$ iff $j = i + 1$ and $V(\mathsf{p}) = \{w_2, w_3\}$, $V(\mathsf{q}) = \{w_1, w_2, w_3, w_4, w_5\}$, $V(\mathsf{r}) = \emptyset$.



1. $\mathfrak{M}, w_1 \models \Diamond \Box \mathsf{p}$
2. $\mathfrak{M}, w_1 \not\models \Diamond \Box \mathsf{p} \rightarrow \mathsf{p}$
3. $\mathfrak{M}, w_2 \models \Diamond (\mathsf{p} \wedge \neg \mathsf{r})$
4. $\mathfrak{M}, w_1 \models \mathsf{q} \wedge \Diamond (\mathsf{q} \wedge \Diamond (\mathsf{q} \wedge \Diamond (\mathsf{q} \wedge \Diamond \mathsf{q})))$
5. $\mathfrak{M} \models \Box \mathsf{q}$

Validity and (Global) Satisfaction

We take on a global point of view.

Given a specification like $\varphi := \Box \neg \text{crash}$. In which states should it be true?

Definition 2.11 (Validity)

A formula φ is called valid or globally true in a model $\mathfrak{M}$ iff $\mathfrak{M}, w \models \varphi$ for all $w \in W_{\mathfrak{M}}$. We write $\mathfrak{M} \models \varphi$.

φ is satisfiable in $\mathfrak{M}$ if $\mathfrak{M}, w \models \varphi$ for some $w \in W_{\mathfrak{M}}$.

Analogously, we say that a set Σ of formulae is valid (resp. satisfiable) in $\mathfrak{M}$ iff all formulae in Σ are valid (resp. satisfiable) in $\mathfrak{M}$.

Validity and satisfiability are dual concepts!

Example 2.12

In which models is the following formula true?

$$\Box\,(\mathsf{p} \rightarrow \mathsf{q}) \rightarrow (\Box\,\mathsf{p} \rightarrow \Box\,\mathsf{q})$$

$$\begin{aligned}
&\mathfrak{M}, w \models \Box\,(\mathsf{p} \rightarrow \mathsf{q})\\
&\text{iff } \forall w'(w\mathcal{R}w' \Rightarrow \mathfrak{M}, w' \models \mathsf{p} \rightarrow \mathsf{q})\\
&\text{iff } \forall w'(w\mathcal{R}w' \Rightarrow (\mathfrak{M}, w' \models \mathsf{p} \Rightarrow \mathfrak{M}, w' \models \mathsf{q}))\\
&\text{implies } \forall w'(w\mathcal{R}w' \Rightarrow \mathfrak{M}, w' \models \mathsf{p}) \Rightarrow\\
&\qquad \forall w'(w\mathcal{R}w' \Rightarrow \mathfrak{M}, w' \models \mathsf{q})\\
&\text{iff } \mathfrak{M}, w \models \Box\,\mathsf{p} \Rightarrow \mathfrak{M}, w \models \Box\,\mathsf{q}\\
&\text{iff } \mathfrak{M}, w \models \Box\,\mathsf{p} \rightarrow \Box\,\mathsf{q}
\end{aligned}$$

The formula is true in any frame and hence in any model.

Modal Consequence Relation

Up to now we verified formulae in a given model and state. Often, it is interesting to know whether a property follows from a given set of formulae.

Definition 2.13 (Local Consequence Relation)

Let $\mathcal{M}$ be a class of models, Σ be a set of formulae and φ be a formula.

- φ is a (local) semantic consequence of Σ over $\mathfrak{M}$, written $\Sigma \models_{\mathcal{M}} \varphi$, if for all $\mathfrak{M} \in \mathcal{M}$ and all $w \in W_{\mathfrak{M}}$ it holds that $\mathfrak{M}, w \models \Sigma$ implies $\mathfrak{M}, w \models \varphi$.
- If $\mathcal{M}$ is the class of all models we just say that φ is a (local) consequence of Σ and write $\Sigma \models \varphi$.

$\rightsquigarrow$ blackboard

Example 2.14

Let $\mathcal{M}$ be the class of transitive models. Then:

1. $\Diamond \Diamond p \models_{\mathcal{M}} \Diamond p$,
2. $\Box p \models_{\mathcal{M}} \Box \Box p$, but
3. $\Box\Box p \models_{\mathcal{M}} \Box p$ does not hold.

Is there a class of models $\mathcal{M}$ for which $\Diamond \Diamond p \models_{\mathcal{M}} \Diamond p$ holds, but no model in $\mathcal{M}$ is transitive?

Definition 2.15 (Global Consequence Relation)

Let $\mathcal{M}$ be a class of models, Σ be a set of formulae and φ be a formula.

- Recall that for $\mathfrak{M} \in \mathcal{M}$, we denote by $\mathfrak{M} \models \Sigma$ that for all $w \in \mathfrak{M} : \mathfrak{M}, w \models \Sigma$.
- φ is a global semantic consequence of Σ over $\mathcal{M}$ (written as $\Sigma \models^g_{\mathcal{M}} \varphi$) if for all $\mathfrak{M} \in \mathcal{M}$ it holds that $\mathfrak{M} \models \Sigma$ implies $\mathfrak{M} \models \varphi$.
- If $\mathcal{M}$ is the set of all models we just say that φ is a global consequence of Σ and write $\Sigma \models^g \varphi$.
- A formula is globally true, or valid in $\mathcal{M}$, if it is true in all models in $\mathcal{M}$.

Example 2.16

Consider the formulae p and $\Box$p and the class $\mathcal{M}$ of all models.

1. p does not locally imply $\Box$p in $\mathcal{M}$.
2. p does globally imply $\Box$p in $\mathcal{M}$.

Note that global validity of a formula ϕ is difficult to achieve:

- ϕ has to be true in all models $\mathfrak{M}$ in $\mathfrak{M}$. And truth in $\mathfrak{M}$ means truth in all $\mathfrak{M}, w$ for all w.

In the next chapter, we will describe such formulae in much more detail.

Frames and Validity

In Example 2.32 we have seen that a formula can be true/false for all valuations. We can speak about structural properties ignoring contingent information.

Definition 2.17 (Frame Validity: $\mathfrak{F} \models \varphi$)

Let $\mathfrak{F}$ be a frame and $\varphi \in \mathcal{L}_{\text{BML}}$.

1. φ is valid in $\mathfrak{F}$ and $w \in W_{\mathfrak{F}}$, written $\mathfrak{F}, w \models \varphi$, if $\mathfrak{M}, w \models \varphi$ for all models $\mathfrak{M} = (\mathfrak{F}, V)$ based on $\mathfrak{F}$.
2. φ is valid in $\mathfrak{F}$, written $\mathfrak{F} \models \varphi$, if $\mathfrak{F}, w \models \varphi$ for all $w \in W_{\mathfrak{F}}$.
3. Let $\mathfrak{K}$ be class of frames. φ is said to be valid in $\mathfrak{K}$, if φ is valid in each frame $\mathfrak{F} \in \mathfrak{K}$.

Example 2.18

Validity differs from truth in many ways:

- When $\varphi \vee \psi$ is true at point w this means: Either φ or ψ is true at w.
- If $\varphi \vee \psi$ is valid on a frame $\mathfrak{F}$ this does **not** mean: Either φ or ψ is valid on $\mathfrak{F}$.
- Counterexample: $q \vee \neg q$.

Note that the last counterexample is similar to the example on Slide 39.

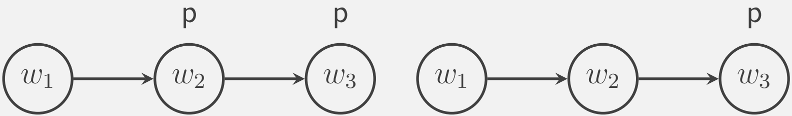
Example 2.19

Is $\Diamond \top$ valid in all frames? In which class is the formula valid?

What about $\Box \top$? $\leadsto$ blackboard

Example 2.20

Is $\Diamond \Diamond p \rightarrow \Diamond p$ true in w_1?



Is there a class of frames in which formula is valid? $\leadsto$

Example 2.21

In which frame is $\Diamond (\mathrm{p} \vee \mathrm{q}) \rightarrow (\Diamond \mathrm{p} \vee \Diamond \mathrm{q})$ valid?
In all Kripke frames!

Let $\mathfrak{F}$ be a frame, $w \in W_{\mathfrak{F}}$, and V be any valuation where $\mathfrak{M} = (\mathfrak{F}, V)$. We have to show that

$$\mathfrak{M}, w \models \Diamond (\mathrm{p} \vee \mathrm{q}) \quad \text{implies} \quad \mathfrak{M}, w \models \Diamond \mathrm{p} \vee \Diamond \mathrm{q}$$

Assume that the antecedent holds. Then, there is a state w' with $\mathcal{R}ww'$ and $\mathfrak{M}, w' \models \mathrm{p} \vee \mathrm{q}$. Hence, we have $\mathfrak{M}, w \models \Diamond \mathrm{p} \vee \Diamond \mathrm{q}$.

Lemma 2.22 (Distribution Axioms)

The two formulae

$$\Diamond (p \lor q) \to (\Diamond p \lor \Diamond q)$$
$$\Box (p \to q) \to (\Box p \to \Box q)$$

are both valid in all Kripke frames $\mathfrak{F}$. *The last formula is also called* axiom K.

Proof.

$\leadsto$ Examples 2.12 and 2.21. $\Box$

2.2 Modal Language $\mathcal{L}_{ML}$

Multiple Relations

We have just introduced two basic modalities:

$\Diamond$ (there exists a successor) and $\Box$ (for all successors).

How many different relations can we describe?

We define a very general logic for talking about relational structures with *multiple relations*.

Multiple Modal Operators

A graph contains multiple relations:

$$\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \ldots, \mathcal{R}_{\text{likes}}, \mathcal{R}_{\text{dislikes}}, \ldots$$

and for each relation we need a diamond operator

$$\langle 1 \rangle, \langle 2 \rangle, \langle 3 \rangle, \ldots, \langle \text{likes} \rangle, \langle \text{dislikes} \rangle, \ldots$$

Sometimes we also write $\langle R_i \rangle$ for $\langle i \rangle$.
Set of modalities: $\mathcal{O}p$

$\langle m \rangle \varphi$: There is a m-sucessor in which φ holds.
$[m]\varphi$: Every m-sucessor satisfies φ.

$$m \in \mathcal{O}p$$

Syntax

Up to now we considered binary relations: $\mathcal{R} \subseteq W \times W$

We generalize this to arbitrary arities. Let $\mathcal{O}p$ be a set of modalities.

Can you think about a sensible ternary relation?

Definition 2.23 (Modal Similarity Type)

A modal similarity type is given by $\tau = (\mathcal{O}p, \rho)$ where $\mathcal{O}p$ is a non-empty set and ρ is a function $\rho : \mathcal{O}p \to \mathbb{N}$.
We also write $m \in \tau$ instead of $m \in \mathcal{O}p$.

Remark 2.24

If the arity is clear from context we omit the function.

Definition 2.25 (Modal Language $\mathcal{L}_{\mathsf{ML}}$)

Let $\tau = (\mathcal{O}p, \rho)$ be a modal similarity type and $\mathcal{P}rop$ be a set of propositions. The **modal language** $\mathcal{L}_{\mathsf{ML}}(\tau, \mathcal{P}rop)$ is given by all formulae defined by the following grammar:

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle m \rangle(\underbrace{\varphi, \ldots, \varphi}_{\rho(m)\text{-times}})$$

where $\mathsf{p} \in \mathcal{P}rop$ and $m \in \mathcal{O}p$.

Definition 2.26 (Macros)

$$[m](\varphi_1, \ldots, \varphi_n) := \neg\langle m \rangle(\neg\varphi_1, \ldots, \neg\varphi_n)$$

$[m]$ is called the **dual** of $\langle m \rangle$ and vice versa.

What can we model with this language?

Definition 2.27 (τ-Kripke Frame)

A τ-(Kripke) frame is given by $\mathfrak{F} = (W, \{\mathcal{R}_m \mid m \in \tau\})$ where

- W is a non-empty set, called set of **domains** or **worlds**,
- $\mathcal{R}_m \subseteq W^{\tau(m)+1}$ are $(\tau(m)$**+1**$)$-ary relations for $m \in \tau$.

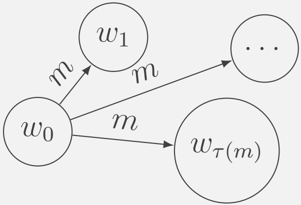
τ is a modal similarity type.
$\tau(m) + 1$ corresponds to the reachable states $\tau(m)$ with $\langle m \rangle$ and the current state (+1).

Frames are mainly used to talk about validities: They stand for a whole set of models.

Kripke frames are simply relational structures!

Example 2.28

$\mathcal{R}_m w_0 \dots w_{\tau(m)}$:



We extend frames by valuations for propositional variables.

Definition 2.29 (τ-Kripke Model)

A τ-(Kripke) model is given by $\mathfrak{M} = (W, \{\mathcal{R}_m \mid m \in \tau\}, V)$ where

- $(W, \{\mathcal{R}_m \mid m \in \mathcal{O}p\})$ is τ-frame
- $V : \mathcal{P}rop \rightarrow \mathcal{P}(W)$ is called labeling function or valuation

Kripke models are simply relational structure with labels.

Remark 2.30

Sometimes we use an analogous definition of valuations: $V : W \rightarrow \mathcal{P}(\mathcal{P}rop)$.

Formal semantics of $\mathcal{L}_{\mathrm{ML}}$.

Definition 2.31 (Semantics $\mathfrak{M}, w \models \varphi$)

Let $\mathfrak{M}$ be a Kripke model, $w \in W_{\mathfrak{M}}$, and $\varphi \in \mathcal{L}_{\mathrm{ML}}$. φ is said to be locally true or satisfied in $\mathfrak{M}$ and world w, written as $\mathfrak{M}, w \models \varphi$, if the following holds:

$\mathfrak{M}, w \models \mathsf{p}$ iff $w \in V_{\mathfrak{M}}(\mathsf{p})$ and $\mathsf{p} \in \mathcal{P}rop$,

$\mathfrak{M}, w \models \neg\varphi$ iff not $\mathfrak{M}, w \models \varphi$

$\mathfrak{M}, w \models \varphi \vee \psi$ iff $\mathfrak{M}, w \models \varphi$ or $\mathfrak{M}, w \models \psi$

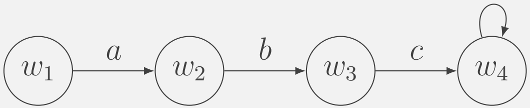
$\mathfrak{M}, w \models \langle m \rangle(\varphi_1, \ldots, \varphi_{\tau(m)})$ iff there are worlds $w_1, \ldots, w_{\tau(m)} \in W$ such that $R_m w w_1 \ldots w_\tau$ and $\mathfrak{M}, w_i \models \varphi_i$ for all $i = 1, \ldots, \tau(m)$

Given a set $\Sigma \subseteq \mathcal{L}_{ML}$ we write $\mathfrak{M}, w \models \Sigma$ iff $\mathfrak{M}, w \models \varphi$ for all $\varphi \in \Sigma$.

Some Examples

Example 2.32

Three unary operators: $\langle a \rangle, \langle b \rangle, \langle c \rangle$
Three relations: $\mathcal{R}_a, \mathcal{R}_b, \mathcal{R}_c$



What can be said about $\langle a \rangle \mathsf{p} \rightarrow \langle b \rangle \mathsf{p}$?

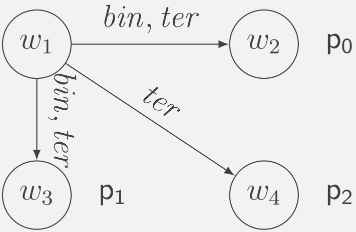
It does not depend on the valuation!

Assume $V(\mathsf{p}) = \{w_2, w_3\}$?
What can be said about $\langle a \rangle \mathsf{p} \rightarrow \langle a \rangle \langle b \rangle \mathsf{p}$?

Example 2.33

Binary operator: $\langle bin \rangle$ Ternary operator: $\langle ter \rangle$

Relations: $\mathcal{R}_{\text{bin}} = \{(w_1, w_2, w_3)\}, \mathcal{R}_{\text{ter}} = \{(w_1, w_2, w_3, w_4)\}$



What about $\langle \text{bin} \rangle(\text{p}_0, \text{p}_1) \to \langle \text{ter} \rangle(\text{p}_0, \text{p}_1, \text{p}_2)$?

Each triangle starting in w is part of a rectangle starting in it.

A Temporal Logic

Reasoning about time:

Example 2.34 (The Basic Temporal Language)

Two unary operators: $\mathcal{O}p = \{F, P\}$.
Modal similarity type: $\tau = (\{F, P\}, \{(F, 1), (P, 1)\})$.

$\langle F\rangle\varphi$: φ will be true at some Future time point.

$\langle P\rangle\varphi$: There is a Past time point, where φ is true.

$[F]\varphi$: φ will always be true.

It is common to use $\Diamond$ for $\langle F\rangle$, $\Box$ for $[F]$, and $\Diamond^{-1}$ for $\langle P\rangle$.

What does the following formula express?

$\Diamond^{-1}\varphi \rightarrow \Box\Diamond^{-1}\varphi$: Whatever has happened will always have happened.

$\rightsquigarrow$ blackboard

Example 2.35 (Bidirectional Frames, Models)

Recall the basic temporal logic. How many relations do we need?

$\mathcal{R}_F$ and $\mathcal{R}_P$ should satisfy the following property:

$$\forall xy(\mathcal{R}_F xy \leftrightarrow \mathcal{R}_P yx)$$

Given one, the other is fixed! We call this kind of frame bidirectional.

What other conditions should be fulfilled to model "time"?
$\rightsquigarrow$ Exercise

A Logic About Programs

Verification of programs:

Example 2.36 (Program π)

$\top \rightarrow \psi$
while true **do**
  if ψ **then**
    $\mu \rightarrow \phi$
    $\psi \rightarrow \mu$
  end if
  if ϕ **then** exit
  end if
end while

Example 2.37 (Propositional Dynamic Logic)

Infinite collection of diamonds: $\mathcal{O}p = \{\pi \mid \pi$ is a program$\}$

What do the following operators express?

$\langle\pi\rangle\varphi$: Some terminating execution of π leads to a state with information φ

$[\pi]\varphi$: Each terminating execution of π leads to a state with information φ

It would be nice if we could combine simple programs to complex ones:

$\pi \cup \pi'$: **Nondeterministic choice**

$\pi ; \pi'$: **Sequential composition**

π^* : **Iterative execution**

What do the following statements express?

$\langle\pi^*\rangle\varphi \leftrightarrow \varphi \vee \langle\pi;\pi^*\rangle\varphi$: A state with information φ is reached by executing π a finite number of times iff the current state satisfies φ or we can execute π once and reach a state in which φ holds by executing π a finite number of times.

$[\pi^*](\varphi \rightarrow [\pi]\varphi) \rightarrow (\varphi \rightarrow [\pi^*]\varphi)$: $\rightsquigarrow$ Exercise.

Do these formulae always hold?

How can we actually use this logic?

Example 2.38 (Regular Frames and Models (1))

How should models for PDL look like? A relation for each program?
We can create new relations by existing ones:

$$\begin{aligned}
\mathcal{R}_{\pi_1 \cup \pi_2} &= \mathcal{R}_{\pi_1} \cup \mathcal{R}_{\pi_2} \\
\mathcal{R}_{\pi_1 ; \pi_2} &= \mathcal{R}_{\pi_1} \circ \mathcal{R}_{\pi_2} \\
\mathcal{R}_{\pi^*} &= (\mathcal{R}_{\pi})^*
\end{aligned}$$

Example 2.39 (Regular Frames and Models (2))

Let Π be the smallest set of programs containing the basic programs and which is closed under $\cup$, $;$, and *. Then a regular frame for Π is a labeled transition system $(W, \{\mathcal{R}_\pi \mid \pi \in \Pi\})$ such that:

- $\mathcal{R}_a$ is an arbitrary binary relation for each basic program a, and
- for all complex programs π, $\mathcal{R}_\pi$ is the binary relation inductively constructed with the rules above.

Syntax & Semantics: Convention

To make life simpler we assume the following:

1. We have one unary modality ($|\mathcal{O}p| = 1$, say $\mathcal{O}p = \{m\}$, with arity 1)
2. We use $\Diamond$ (resp. $\Box$) for $\langle m \rangle$ (resp. $[m]$).
3. We do not explicitly mention the modal similarity type and drop τ from all Definitions, if it is clear from context.
4. We use $\mathcal{R}$ for the binary relation between worlds. We use both notations $\mathcal{R}w_1w_2$ and $w_1\mathcal{R}w_2$ to denote that world w_2 is reachable from world w_1.

2.3 Extending ML to FOL

Modal vs. First-Order Language?

What are the main differences between the languages?

First-order logic

1. contains constants.
2. contains predicate and function symbols.
3. allows global quantification.

On the other hand, modal logic

1. has modal operators.
2. is only local.

How can we extend our current modal language to FOL?

Modal Language vs. First-Order Logic

Function symbols and k-ary relation symbols for $k > 2$ can be encoded as relations as shown in the previous chapter. Therefore, we restrict ourselves to FOL without function symbols but constants.

The main idea is that we identify elements from the first-order domain (constants) with worlds in the Kripke model.

In the following we introduce **syntactic means** to refer to these worlds.

Basic Hybrid Language

Modal logic does not allow to name worlds. We cannot say that **some formula holds at a particular world**.

For example, we cannot say that a specific state is blue and black.

We introduce the basic hybrid language which has the described feature.

What does a unique name for a point correspond to?

Constants!

We add a second sort of propositional symbols to our modal language.

The new symbols $\mathcal{N}om = \{i, j, k, l, \dots\}$ are called nominals and are used as labels for points. We require that they are disjunct from the propositional symbols, i.e. $\mathcal{N}om \cap \mathcal{P}rop = \emptyset$. In the models nominals are interpreted as singleton sets, i.e. they are true at exactly one world.

In the language we treat nominals as ordinary propositions!

What can you say about the validity of the following formulae?

1. $\Diamond(\mathsf{r} \wedge \mathsf{p}) \wedge \Diamond(\mathsf{r} \wedge \mathsf{q}) \rightarrow \Diamond(\mathsf{p} \wedge \mathsf{q})$
2. $\Diamond(i \wedge \mathsf{p}) \wedge \Diamond(i \wedge \mathsf{q}) \rightarrow \Diamond(\mathsf{p} \wedge \mathsf{q})$

What do we have to ensure with respect to the labeling of nodes by nominals?

Each nominal has to be true at exactly one world!

Given a Kripke model $\mathfrak{M}$ we call $V(i)$ the denotation of i where $i \in \mathcal{N}om$.

In order to capture FOL, we must be able to "jump to a specific world". Therefore we use the satisfaction operators $@_i$:

$@_i\varphi$ iff φ holds at the world labeled with i.

Syntax of the hybrid language

Definition 2.40 (Basic Hybrid Language $\mathcal{L}_{\text{BHL}}$)

Let $\mathcal{P}rop$ be a set of propositions, $\mathcal{N}om$ be a set of nominals, and $\tau = (\mathcal{O}p, \rho)$ be a modal similarity type. The basic hybrid language $\mathcal{L}_{\text{BHL}}(\tau, \mathcal{P}rop, \mathcal{N}om)$ is given by all formulae defined by the following grammar:

$$\varphi ::= \mathsf{p} \mid i \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle m\rangle(\underbrace{\varphi, \ldots, \varphi}_{\rho(m)\text{-times}}) \mid @_i\varphi$$

where $m \in \tau$, $\mathsf{p} \in \mathcal{P}rop$ and $i \in \mathcal{N}om$.

Semantics

We say that a Kripke model is hybrid if, and only if, $V : \mathcal{Prop} \cup \mathcal{Nom} \to \mathcal{P}(W)$ and each nominal is true at exactly one world, i.e. $|V(i)| = 1$ for all $i \in \mathcal{Nom}$.

In the following, when considering the hybrid language we implicitly assume that the models are hybrid.

Definition 2.41 (Semantics of $\mathcal{L}_{\text{BHL}}$)

Let $\mathfrak{M}$ be a (hybrid) Kripke model, $w \in W_{\mathfrak{M}}$, $\mathcal{Nom}$ a set of nominals and $\varphi \in \mathcal{L}_{\text{BHL}}$.
The semantics of the basic hybrid language extends the semantics of modal logic by the following two clauses:

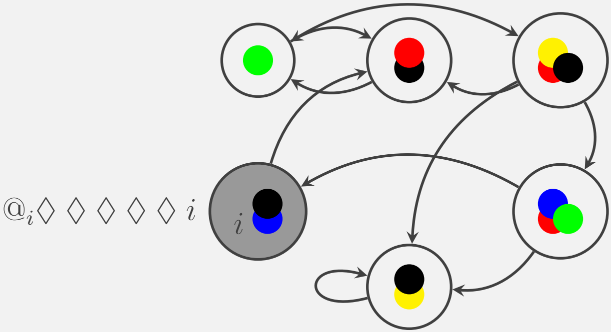
$\mathfrak{M}, w \models i$ iff $V(i) = \{w\}$ and $i \in \mathcal{Nom}$

$\mathfrak{M}, w \models @_i\varphi$ iff $\mathfrak{M}, w' \models \varphi$ where $V(i) = \{w'\}$

Comparing Nodes

Example 2.42 (Colored Graph II)

How to say: “We can come back to the same node”?



Example 2.43

In ML we have seen how to describe **frame properties** by **modal formulae**. It turns out that the hybrid language is much more expressive in that respect. Consider a frame $(W, \mathcal{R})$ and give the corresponding property of R to each of the following formulae:

1. $i \rightarrow \Diamond\, i$ (reflexivity)
2. $\Diamond\, \Diamond\, i \rightarrow \Diamond\, i$ (transitivity)
3. $\Diamond\, i \rightarrow \Diamond\, \Diamond\, i$ (density)
4. $i \rightarrow \neg\Diamond\, i$ (irreflexivity)
5. $\Diamond\, \Diamond\, i \rightarrow \neg\Diamond\, i$ (intransitivity)

Simulating FOL: 1. Attempt

We construct the following two translations

TR: Transformation of FOL models to Kripke models

tr: Transformation of FOL formulas to Modal formulas

such that for a *given* FOL model $\mathfrak{M}_{\mathsf{FOL}}$ and *all* FOL formulas φ it holds that

$$\mathfrak{M}_{\mathsf{FOL}} \models \varphi \qquad \text{iff} \qquad TR(\mathfrak{M}_{\mathsf{FOL}}), w \models tr(\varphi)$$

Translations of Models

Let $\mathfrak{M} = (U, I)$ be a FOL model over $\mathcal{P}red$ without function symbols (constants are allowed).

We construct $\mathfrak{M}' := TR(\mathfrak{M}) = (W, \mathcal{R}, V)$ as follows:

1. $W := U \cup \{w_0\}$ where w_0 is a new world not occurring in U.
2. For each $c \in \mathcal{F}unc^0 \cup \mathcal{V}ar$ label the world $I(c)$ with c, i.e. $V(c) := \{I(c)\}$. We set $\mathcal{N}om = \mathcal{F}unc^0 \cup \mathcal{V}ar$.
3. For each $P \in \mathcal{P}red^1$ we label each world $w \in \mathfrak{M}(P)$ with P and for $\mathsf{p} \in \mathcal{P}red^0$ label w_0 with p.
4. For each $P \in \mathcal{P}red^k$ with $k > 1$ we add a k-ary relation R_P to $\mathcal{R}$ (resp. to $\mathcal{O}p$) such that $R_P w_1 w_2 \dots w_k$ iff $(w_1, w_2, \dots, w_k) \in \mathfrak{M}(P)$.

We shall also identify R_P with P if no confusion occurs.

Example 2.44

Consider the follwing FOL model: $\mathfrak{M} = (\{1, 2, 3\}, I)$ over $\mathcal{F}unc^0 = \{a, b\}$, $\mathcal{V}ar = \{x\}$, $\mathcal{P}red^1 = \{P^1\}$, and $\mathcal{P}red^2 = \{P^2\}$ with

1. $I(P^1) = \{1, 3\}$
2. $I(P^2) = \{(1, 2), (1, 3)\}$
3. $I(a) = 2$
4. $I(b) = 1$
5. $I(x) = 2.$

Is the following formulae true in $\mathfrak{M}$?

$$\exists x P(b, x) \wedge \forall y \neg P(x, b)$$

How does the Kripke model look like? $\rightsquigarrow$ blackboard

Translation $tr : \mathcal{L}_{FOL} \to \mathcal{L}_{\mathsf{BHL}}$

Equality: $tr(s \doteq t) := @_s t$

Unary predicates: $tr(P(x)) := @_x P$

Predicates: $tr(R(x_1, \ldots, x_k)) := @_{x_1} \langle R \rangle (x_2, \ldots, x_k)$

Negation: $tr(\neg\varphi) := \neg tr(\varphi)$

Or: $tr(\varphi \vee \psi) := tr(\varphi) \vee tr(\psi)$

Quantification: $tr(\exists x \varphi) := ?$

We cannot yet say: There exists a world such that...!

Renaming worlds

Recall the semantics of FOL:

$$\mathfrak{M} \models \exists x \varphi \quad \texttt{iff} \quad \mathfrak{M}^{[x/a]} \models \varphi$$

for some $a \in U$ where $\mathfrak{M}^{[x/a]}$ denotes the model equal to $\mathfrak{M}$ but $\mathfrak{M}^{[x/a]}(x) = a$

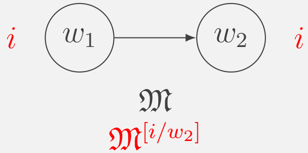
We need a model update similar to the first-order case.

Definition 2.45 (Model Update $\mathfrak{M}^{[i/w]}$)

Let $\mathfrak{M}$ be a hybrid Kripke model, $w \in W_{\mathfrak{M}}$, and $i \in \mathcal{N}om$
The define $\mathfrak{M}^{[i/w]}$ to be the model equal to $\mathfrak{M}$ but the *denotation* of i is set to w, i.e. $V_{\mathfrak{M}^{[i/w]}}(i) = \{w\}$.

Note that we use the same notation as for FOL.

Example 2.46



We still lack a syntactic counterpart!

Extended syntax

Definition 2.47 (HL With Renaming $\mathcal{L}_{\mathsf{HLR}}$)

Let $\tau = (\mathcal{O}p, \rho)$ be a modal similarity type and $\mathcal{P}rop$ be a set of propositions, and $\mathcal{N}om$ be a set of nominals. The **hybrid language with renaming** $\mathcal{L}_{\mathsf{HLR}}(\tau, \mathcal{P}rop, \mathcal{N}om)$ is given by all formulae defined by the following grammar:

$$\varphi ::= i \mid \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle m\rangle(\underbrace{\varphi, \ldots, \varphi}_{\rho(m)\text{-times}}) \mid @_i\varphi \mid r_i\varphi$$

where $m \in \tau$, $\mathsf{p} \in \mathcal{P}rop$ and $i \in \mathcal{N}om$.

Semantics

Definition 2.48 (Semantics of $\mathcal{L}_{\text{HLR}}$)

Let $\mathfrak{M}$ be a (hybrid) Kripke model, $w \in W_{\mathfrak{M}}$ and $\varphi \in \mathcal{L}_{\text{BHL}}$. The semantics of the basic hybrid language with renaming extends the semantics of the hybrid language by the following clause:

$\mathfrak{M}, w \models r_i\varphi$ iff $\mathfrak{M}^{[i/w']}, w \models \varphi$ for some $w' \in W$.

Consider the analogy: $\exists x \varphi(x)$ vs. $r_i\varphi(i)$

Simulating first-order logic

Definition 2.49 ($tr : \mathcal{L}_{FOL} \to \mathcal{L}_{\mathsf{HLR}}$)

Prop. symbols: $tr(\mathsf{p}) := @_{w_0}\mathsf{p}$

Equality: $tr(s \doteq t) := @_s t$

Unary predicates: $tr(P(x)) := @_x P$

Predicates: $tr(R(x_1, \ldots, x_k)) := @_{x_1}\langle R\rangle(x_2, \ldots, x_k)$

Negation: $tr(\neg\varphi) := \neg tr(\varphi)$

Or: $tr(\varphi \vee \psi) := tr(\varphi) \vee tr(\psi)$

Quantification: $tr(\exists x\varphi) := r_x tr(\varphi)$

Example 2.50

Let $\varphi := \exists x(banker(x) \rightarrow \exists y(customer(y) \wedge spoofs(x,y)))$. We translate φ to a ML formula.

- $tr(\exists x(banker(x) \rightarrow \exists y(customer(y) \wedge spoofs(x,y))))$
- $r_x(tr(banker(x) \rightarrow \exists y(customer(y) \wedge spoofs(x,y))))$
- $r_x((tr(banker(x)) \rightarrow tr(\exists y(customer(y) \wedge spoofs(x,y)))))$
- $r_x(@_x banker \rightarrow tr(\exists y(customer(y) \wedge spoofs(x,y))))$
- $r_x(@_x banker \rightarrow r_y(tr(customer(y) \wedge spoofs(x,y))))$
- $r_x(@_x banker \rightarrow r_y(tr(customer(y)) \wedge tr(spoofs(x,y))))$
- $r_x(@_x banker \rightarrow r_y(@_y customer \wedge @_x \langle spoofs \rangle y))$

Example 2.51

Consider the model and formula from Example 2.44: $\mathfrak{M} = (\{1,2,3\}, I)$ over $\mathcal{F}unc^0 = \{a, b\}$, $\mathcal{V}ar = \{x\}$, $\mathcal{P}red^1 = \{P^1\}$, and $\mathcal{P}red^2 = \{P^2\}$ with

1. $I(P^1) = \{1,3\}$
2. $I(P^2) = \{(1,2),(1,3)\}$
3. $I(a) = 2$
4. $I(b) = 1$
5. $I(x) = 2.$

Construct $tr(\varphi)$ where

$$\varphi = \exists x P(b,x) \land \forall y \neg P(x,b)$$

and show that $TR(\mathfrak{M}), w_0 \models tr(\varphi)$. $\rightsquigarrow$ blackboard

Theorem 2.52

Let $\mathcal{F}unc$ *contain only constants and let* $\varphi \in \mathcal{L}_{FOL}(\mathcal{V}ar, \mathcal{P}red, \mathcal{F}unc)$. *Then, we have*

$$\mathfrak{M}_{\mathsf{FOL}} \models \varphi \quad \textit{iff} \quad TR(\mathfrak{M}_{\mathsf{FOL}}), w_0 \models tr(\varphi).$$

Proof.

By structural induction. $\rightsquigarrow$ Exercise □

Translation: The Other Direction

Have a closer look at the semantic clause of the hybrid modal language:

$$\mathfrak{M}, w \models \langle R \rangle \varphi \text{ iff there is an } w' \in W_{\mathfrak{M}} \text{ s.t. } Rww' \text{ and } \mathfrak{M}, w' \models \varphi$$

In which formal language can this be expressed?

In FOL! The translation is immediate:

$$tr_w(\langle R \rangle \varphi) = \exists w' \; (R(w, w') \land tr_{w'}(\varphi))$$

We need w' in $tr_{w'}$ to keep track of the current world.

Translating models

Let $\mathfrak{M} = (W, \{\mathcal{R}_i \mid i \in \tau\}, V)$ be a τ-**Kripke model**. We construct the following **FOL model** $\mathfrak{M}' = (U, I)$:

1. $U = W$
2. $\mathcal{P}red^1 := \mathcal{P}rop$
3. $\mathcal{V}ar := \mathcal{N}om \cup \{x_1, x_2, \dots\}$
4. $\mathcal{F}unc^0 := W$
5. $I(p) := V(p)$ for $p \in \mathcal{P}rop$
6. $I(i) := w$ with $V(i) = \{w\}$ for $i \in \mathcal{N}om$
7. $I(w) := w$
8. $\mathcal{P}red^k := \{P_R \mid R$ is a k-ary accessibility relation$\}$ ($k \geq 2$)
9. $I(P_R) := R$ for $P_R \in \mathcal{P}red^k$

Translating models

Example 2.53

Let $\mathcal{P}rop = \{\mathsf{p}\}$ and $\mathfrak{M} = (\{w_1, w_2, w_3\}, \mathcal{R}, V)$ be a Kripke model with $\mathcal{R} = \{(w_2, w_3)\}$ and $V(\mathsf{p}) = \{w_3\}$.

Which of the the following hold?

1. $\mathfrak{M}, w_1 \models r_i\langle R\rangle\mathsf{p}$
2. $\mathfrak{M}, w_1 \models r_i@_i\langle R\rangle\mathsf{p}$

Construct the FOL-model $TR(\mathfrak{M})$. $\rightsquigarrow$ blackboard

Translating formulae

Definition 2.54 ($tr_x : \mathcal{L}_{\textsf{HLR}} \to \mathcal{L}_{FOL}$)

We define the translation recursively as follows:

Atomic Propositions: $tr_x(P) := P(x)$

Nominals: $tr_x(i) := (x \doteq i)$

Negation: $tr_x(\neg\varphi) := \neg tr_x(\varphi)$

Disjunction: $tr_x(\varphi \lor \psi) := tr_x(\varphi) \lor tr_x(\psi)$

Modality: $tr_x(\langle R\rangle(\varphi_1, \ldots, \varphi_k)) :=$
$\exists y_1 \ldots \exists y_k (P_R(x, y_1 \ldots y_k) \land \bigwedge_{y_i} tr_{y_i}(\varphi_i)$

Satisfaction operator: $tr_x(@_i\varphi) := tr_i(\varphi)$

Renaming: $tr_x(r_i\varphi) := \exists i(tr_x(\varphi))$

Translating models

Example 2.55

Consider the model and formulae from Example 2.53:

1. $\mathfrak{M}, w_1 \models r_i \langle R \rangle \mathsf{p}$
2. $\mathfrak{M}, w_1 \models r_i @_i \langle R \rangle \mathsf{p}$

Translate the formulae to FOL and check whether they are true in the FOL-model $TR(\mathfrak{M})$. $\rightsquigarrow$ blackboard

Theorem 2.56

Let $\varphi \in \mathcal{L}_{HLR}(\tau, \mathcal{P}rop, \mathcal{N}om)$ *and* $\mathfrak{M}$ *be a compatible* τ*-Kripke model. Then, we have*

$$\mathfrak{M}, w \models \varphi \quad \textit{iff} \quad TR(\mathfrak{M}) \models tr_w(\varphi).$$

Proof.

By structural induction $\rightsquigarrow$ Exercise □

2.4 The Standard Translation

The function $tr_x : \mathcal{L}_{\text{HLR}} \to \mathcal{L}_{FOL}$ is a mapping from $\mathcal{L}_{\text{HLR}}$ to $\mathcal{L}_{FOL}$ given a world w of the Kripke model. Here, we would like to define the correspondent language of the (basic) modal logic.

We would like to identify to which sublanguage of first-order logic the (basic) modal logic corresponds to. Of course, this should be a strict subclass of $\mathcal{L}_{FOL}$.

Definition 2.57 (Correspondence Language $\mathcal{L}^1_\tau$)

Let $\mathcal{L}_{\mathsf{ML}}(\tau, \mathcal{P}rop)$ be given. The first-order language corresponding to $\mathcal{L}_{\mathsf{ML}}(\tau, \mathcal{P}rop)$, $\mathcal{L}^1_\tau(\mathcal{P}rop)$, is built over the unary prediactes

$$\mathcal{P}red^1 = \{P_i \mid p_i \in \mathcal{P}rop\}$$

and the following $n+1$ ary predicates

$$\mathcal{P}red^{n+1} = \{P_m \mid m \in \tau \text{ and } \tau(m) = n\}.$$

That is, for each n-ary modal operator m a $n+1$ ary predicate symbol is defined.
$\varphi(x)$ denotes a first-order formula with one free variable, that is, a variable which is not bound by any quantifier.

We define the standard translation ST_x as the restriction of tr_x to $\mathcal{L}_{\mathsf{ML}}(\tau, \mathcal{P}rop)$.

Definition 2.58 (Standard Translation ST_x)

Let x be a FOL variable. The standard translation

$$ST_x : \mathcal{L}_{\mathsf{ML}}(\tau, \mathcal{P}rop) \to \mathcal{L}^1_\tau(\mathcal{P}rop)$$

is defined as follows:

$$\begin{aligned}
ST_x(\mathsf{p}) &= P(x) \\
ST_x(\neg\varphi) &= \neg ST_x(\varphi) \\
ST_x(\varphi \vee \psi) &= ST_x(\varphi) \vee ST_x(\psi) \\
ST_x(\langle m \rangle(\varphi_1, \dots, \varphi_n)) &= \exists y_1 \dots \exists y_n (P_m(x, y_1, \dots, y_n) \wedge \\
&\qquad ST_{y_1}(\varphi_1) \wedge \dots \wedge ST_{y_n}(\varphi_n))
\end{aligned}$$

Theorem 2.59 (Standard Translation = ML)

Let $\mathfrak{M}$ be a τ-model and $\bar{\mathfrak{M}}$ the FOL interpretation of $\mathfrak{M}$. Then, the following holds

1. $\mathfrak{M}, w \models \varphi$ *iff* $\bar{\mathfrak{M}}^{[x/w]} \models ST_x(\varphi)$
2. $\mathfrak{M} \models \varphi$ *iff* $\bar{\mathfrak{M}} \models \forall x ST_x(\varphi)$

Proof: $\rightsquigarrow$ Exercise

3. Basic Modal Logics

3 Basic Modal Logics

- Inferences and Properties
- Normal Modal Logics
- Sound- and Completeness
- Finite Models via Filtration

Content of this Chapter (1)

In this chapter we introduce the most important modal logics: normal modal logics. We define them, by giving appropriate axiom systems. To this end, we have to introduce proofs in modal logics (Definition 3.13). Then we link these axioms to properties of the transition relation between worlds. It is much easier to work with these semantical notions, than with proof-theoretical means.

Content of this Chapter (2)

After stating and discussing the implications of the sound- and completeness results, we formally prove them using the idea of canonical models: **Any modal logic is strongly complete wrt its canonical model**. Most completeness results for modal logics (wrt. a particular class of frames) are instances of this theorem. We end this chapter with a discussion of the finite model property and a method, the filtration theorem, to obtain finite models. Using this method, many modal logics can be shown to be decidable.

3.1 Inferences and Properties

The Story So Far . . .

We have introduced several languages: e.g, the

- propositional language,
- basic modal language,
- modal language,
- first-order language.

Truth of formulae in these languages was defined by **models**: in a **semantic** way. Using this notion, we **purely semantically** defined when a formula follows from a set of formulae:

$$\Sigma \models \varphi \text{ iff for all models } \mathfrak{M}\text{: } \mathfrak{M} \models \Sigma \Rightarrow \mathfrak{M} \models \varphi$$

In modal logic, we also distinguished between **local** and global consequences: Truth in $\mathfrak{M}, w$ and truth in $\mathfrak{M}$.

Syntactic Inferences

First Question

Given a class $\mathcal{F}$ of frames (or models). Which formulae are valid in all these frames?

Note that there might be classes of frames, where this problem is undecidable: We can not find an algorithm to determine whether a given formula is valid or not.

Second Question

Can we define a calculus (like the **Tableaux-method**) to systematically produce all valid formulae?

We will indeed define several systems of axioms, which produce exactly the valid formulae in particular classes of frames (see Theorem 3.29).

Valid Formulae in an Arbitrary Frame

Which formulae are true in the class of all frames $\mathcal{F}$?

Axioms: **Which formulae from propositional logic are valid in it?**
Certainly the **tautologies** of propositional logic.
And the axiom **K**:

$$\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$$

Closure: What about **Modus Ponens**?
What about substitution? We know that $p \vee \neg p$ is a tautology. Then so should $\phi \vee \neg\phi$ be, where ϕ is **any modal logic formula**.

Necessitation: Suppose the formula ϕ is true in the frame $\mathcal{F}$. Then so is $\Box\phi$.

We take a closer look at these (inference) rules soon.

Validity in a Particular Frame (1)

Suppose we do know something about a particular frame.

Serial: Frames where the relation $\mathcal{R}$ is **serial**:
For all w there is a w' with $w\mathcal{R}w'$.

Reflexive: Frames where the relation $\mathcal{R}$ is **reflexive**:
For all w: $w\mathcal{R}w$.

Transitive: Frames where the relation $\mathcal{R}$ is **transitive**:
For all w, w', w'': $w\mathcal{R}w'$ and $w'\mathcal{R}w''$ implies $w\mathcal{R}w''$.

Euclidean: Frames where the relation $\mathcal{R}$ is **euclidean**:
For all w, w', w'': $w\mathcal{R}w'$ and $w\mathcal{R}w''$ implies $w'\mathcal{R}w''$.

Symmetric: Frames where the relation $\mathcal{R}$ is **symmetric**:
For all w, w': $w\mathcal{R}w'$ implies $w'\mathcal{R}w$.

Remember Slide 127 where we discussed transitivity.

Validity in a Particular Frame (2)

Definition 3.1 (Important Formulae)

We assume that we have available one modality operator **O** (corresponding to $\Box$). Later, we consider logics with several operators each of which can satisfy certain properties. When considering only one modality, we write $\Box$ instead of **O**.

K	$\mathbf{O}(p \rightarrow q) \rightarrow (\mathbf{O}p \rightarrow \mathbf{O}q)$
D	$\neg\mathbf{O}(p \wedge \neg p)$
T	$\mathbf{O}p \rightarrow p$
4	$\mathbf{O}p \rightarrow \mathbf{O}\mathbf{O}p$
5	$\neg\mathbf{O}p \rightarrow \mathbf{O}\neg\mathbf{O}p$
B	$p \rightarrow \mathbf{O}\neg\mathbf{O}\neg p$

Validity in Several Frames (3)

Lemma 3.2 (Appropriate Frames)

Let $(W, \mathcal{R})$ *be a Kripke frame. Then the following holds:*

K: $(W, \mathcal{R}) \models \Box(p \to q) \to (\Box p \to \Box q)$.

D: $(W, \mathcal{R}) \models \neg\Box(p \wedge \neg p)$ *iff* $\mathcal{R}$ *is* serial.

T: $(W, \mathcal{R}) \models \Box p \to p$ *iff* $\mathcal{R}$ *is* reflexive.

4: $(W, \mathcal{R}) \models \Box p \to \Box\Box p$ *iff* $\mathcal{R}$ *is* transitive.

5: $(W, \mathcal{R}) \models \neg\Box p \to \Box\neg\Box p$ *iff* $\mathcal{R}$ *is* euclidean.

B: $(W, \mathcal{R}) \models p \to \Box\Diamond p$ *iff* $\mathcal{R}$ *is* symmetric.

Validity in Several Frames (4)

Proof: (1).

The first fact has been proved in Lemma 2.22.

We identify in Kripke models $(W, \mathcal{R}, V)$ **the mapping** V **with subsets** S **of** W**:** $V(\mathsf{p}) = S$ (note that we only need to consider V that make p true or not).

Then $(W, \mathcal{R}) \models \neg\Box\,(\mathsf{p} \wedge \neg\mathsf{p})$ iff

$$\forall S \subseteq W \forall w \in W\ (\exists x (w \mathcal{R} x\ \wedge\ (x \in S \vee x \notin S)))$$

□

Validity in Several Frames (5)

Proof: (2).

$(W, \mathcal{R}) \models \Box \mathsf{p} \rightarrow \mathsf{p}$ **iff** $\forall S \subseteq W \forall w \in W$:

$$\forall x (w \mathcal{R} x \rightarrow x \in S) \rightarrow w \in S.$$

$(W, \mathcal{R}) \models \Box \mathsf{p} \rightarrow \Box \Box \mathsf{p}$ **iff** $\forall S \subseteq W \forall w \in W$:

$$\forall x (w \mathcal{R} x \rightarrow x \in S) \rightarrow \forall x (w \mathcal{R} x \rightarrow \forall y (x \mathcal{R} y \rightarrow y \in S)).$$

$(W, \mathcal{R}) \models \neg \Box \mathsf{p} \rightarrow \Box \neg \Box \mathsf{p}$ **iff** $\forall S \subseteq W \forall w \in W$:

$$\exists x (w \mathcal{R} x \wedge x \notin S) \rightarrow \forall x (w \mathcal{R} x \rightarrow \exists y (x \mathcal{R} y \wedge y \notin S)).$$

$(W, \mathcal{R}) \models \mathsf{p} \rightarrow \Box \Diamond \mathsf{p}$ **iff** $\forall S \subseteq W \forall w \in W$:

$$(w \in S) \rightarrow \forall x (w \mathcal{R} x \rightarrow \exists y (x \mathcal{R} y \wedge y \in S)).$$

Validity in Several Frames (6)

Proof: (3).

The proof is now simple. The "$\Leftarrow$" is trivial. The other direction is by **choosing the right Kripke model** (resp. the right set $S \subseteq W$). For reflexivity and transitivity, we choose $S := \{x : w\mathcal{R}x\}$. For euclidean, we set $S := \{w''\}$. For symmetry we set $S := \{w\}$. □

Recall that

- a binary relation $\mathcal{R}$ is irreflexive, if there is no x with $x\mathcal{R}x$,
- a preorder $\mathcal{R}$ is a binary relation that is **reflexive**, and **transitive**,
- a partial order $\mathcal{R}$ is a **preorder** that is **antisymmetric** (i.e. $\forall x \forall y((x\mathcal{R}y \wedge y\mathcal{R}x) \rightarrow x = y)$),
- a finite partial order $\mathcal{R}$ is a partial order defined on a finite set,
- an equivalence relation $\mathcal{R}$ is a **preorder** that is **symmetric** (i.e. $\forall x \forall y(x\mathcal{R}y \rightarrow y\mathcal{R}x)$).

Lemma 3.3 (Characterising Equivalences)

A binary relation is an equivalence relation if and only if it is reflexive and euclidean.

$\rightsquigarrow$ Exercise

Preserving Truth/Validity

Definition 3.4 (Preserving Frame Validity)

An inference rule $\frac{\varphi_1,\dots\varphi_n}{\psi}$ is called frame validity preserving, if given a frame $\mathcal{F}$ with $\mathcal{F} \models \varphi_1 \wedge \ldots \wedge \varphi_n$, then $\mathcal{F} \models \psi$ as well.

Similarly, a rule is called locally truth preserving for a class of models $\mathcal{M}$, if for $\mathfrak{M} \in \mathcal{M}$ and all $w \in W_{\mathfrak{M}}$ it holds that $\mathfrak{M}, w \models \varphi_1 \wedge \ldots \wedge \varphi_n$ implies $\mathfrak{M}, w \models \psi$.

Analogously, a rule is called globally truth preserving for a class of models $\mathcal{M}$, if for $\mathfrak{M} \in \mathcal{M}$ it holds that $\mathfrak{M} \models \varphi_1 \wedge \ldots \wedge \varphi_n$ implies $\mathfrak{M} \models \psi$.

Modus Ponens

This rule was already introduced on Slide 41.

Definition 3.5 (Modus Ponens)

Let $\varphi, \psi \in \mathcal{L}_{\mathsf{ML}}$. Modus ponens denotes the inference rule which allows to derive ψ from φ and $\varphi \rightarrow \psi$:

$$\frac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$$

Example 3.6

We consider the set $T = \{\Box p, \Diamond \neg p \vee \Diamond q\}$. At first sight, Modus Ponens does not seem to be applicable. However

$$\Diamond \neg p \vee \Diamond q \text{ is equivalent to } \Box p \rightarrow \Diamond q$$

Therefore we can derive from T the formula $\Diamond q$.

Note that if we look at it from a **semantical perspective**, the derivation is trivial: If p is true in **all reachable worlds**, then it is also true in **at least one reachable world**.

Lemma 3.7 (MP is Locally Truth Preserving)

The Modus Ponens rule is locally truth preserving *and therefore also* preserving frame validity*.*

Generalization (Necessitation)

There is another important rule: The generalization or necessitation rule.

Definition 3.8 (Necessitation (N))

The necessitation rule says that whenever φ is valid then so is $\Box\varphi$:

$$\frac{\varphi}{\Box\varphi}$$

It is easy to see that necessitation preserves frame validity. But differently from the modus ponens rule, it is not locally truth preserving!

$\rightsquigarrow$ Exercise: Prove that the rule is frame validity preserving, but not locally truth preserving!

A first Application of Necessitation

We are using the necessitation rule to **prove** the following

Lemma 3.9 (An (In-)Formal Proof)

The formula $\Box(\phi \wedge \phi') \rightarrow (\Box\phi \wedge \Box\phi')$ *is valid in all frames.*

Proof.

We could show this directly, as in Lemma 2.22. However, in view of the formal calculi to come, we choose a more syntactical method.

1. $(\phi \land \phi') \to \phi$ and $(\phi \land \phi') \to \phi'$ are propositional tautologies.
2. Using necessitation, we have $\Box\,((\phi \land \phi') \to \phi)$ and $\Box\,((\phi \land \phi') \to \phi')$.
3. Using the Distribution axiom from Lemma 2.22, we get $\Box\,(\phi \land \phi') \to \Box\,\phi$ and $\Box\,(\phi \land \phi') \to \Box\,\phi'$.
4. The last step follows by using the propositional calculus: If we have ψ and ψ' then we also have $\psi \land \psi'$.

$\Box$

Inference Rules versus Implications

What about the following deduction?

What Went Wrong?

Let us assume that p is true (although it is obviously not valid on all frames). Then we use necessitation and obtain $\Box p$. So, we have proven that $p \rightarrow \Box p$!

Uniform Substitution

The last method to obtain new formulae from old ones is called uniform substitution. This is the formalization of formula schemata.

Let us consider the tautology $p \rightarrow \neg p$. It is easy to see that it does not depend on the particular proposition p. We should be able to substitute any formula for p. The notion of uniform substitution formalizes this observation.

Definition 3.10 (Uniform Substitution)

A **basic substitution** is a mapping $\sigma : \mathcal{P}rop \to \mathcal{L}_{ML}(\mathcal{P}rop)$. A basic substitution *induces* a **(uniform) substitution**

$$(\cdot)\sigma^* : \mathcal{L}_{ML}(\mathcal{P}rop) \to \mathcal{L}_{ML}(\mathcal{P}rop)$$

over formulae as follows:

$$\begin{aligned}
(\bot)\sigma^* &:= \bot \\
(p)\sigma^* &:= \sigma(p) \\
(\neg\varphi)\sigma^* &:= \neg(\varphi)\sigma^* \\
(\varphi \vee \psi)\sigma^* &:= (\varphi)\sigma^* \vee (\psi)\sigma^* \\
(\Diamond\, \varphi)\sigma^* &:= \Diamond\, (\varphi)\sigma^*
\end{aligned}$$

We identify σ^* with σ and omit the parentheses.

Definition 3.11 (Substitution Rule (Subst))

We say that ψ is a substitution instance of φ if there is some uniform substitution σ^* such that $\psi = \varphi\sigma^*$.
The uniform substitution inference rule is as follows:

$$\frac{\varphi}{\varphi\sigma^*}$$

where σ^* is some uniform substitution.

Example 3.12

The formula $\Box\varphi \vee \neg\Box\varphi$ is a substitution instance of $\mathrm{p} \vee \neg\mathrm{p}$.

And so is the formula $\Box\Diamond\neg\varphi \vee \neg\Box\Diamond\neg\varphi$.

Provability

Definition 3.13 (Σ-Proof, Σ)

Let Σ be a set of modal formulae and *Rul* a set of inference rules. A Σ^{Rul}-proof for φ is a finite sequence of formulae $\varphi_1, \ldots, \varphi_n$ s.t.

1. $\varphi = \varphi_n$, and
2. φ_i is either from Σ, or a propositional tautology or obtained by applying the rules in *Rul* to some formulae φ_j $(j < i)$.

We consider $Rul = \{\mathbf{MP}, \mathbf{Subst}\}$ and, for normal modal logics, $Rul = \{\mathbf{MP}, \mathbf{Subst}, \mathbf{N}\}$. When clear from context, we write $\mathbf{\Sigma}$ for the calculus based on *Rul*. Formulae $\varphi \in \Sigma$ are called axioms of $\mathbf{\Sigma}$. We write $\vdash_{\mathbf{\Sigma}} \varphi$. For a set Γ of formulae, we also write $\Gamma \vdash_{\mathbf{\Sigma}} \varphi$ for $\vdash_{\mathbf{\Sigma} \cup \Gamma} \varphi$: φ can be proved from Γ in the system $\mathbf{\Sigma}$.

We have already seen an **informal** proof of the statement in Lemma 3.9. The reverse implication is also true and we give a formal proof of it. The only axiom we need is the formula **K** introduced on Slide 194.

Lemma 3.14 (Consequence of Axiom K)

$$\vdash_{\{K\}} (\Box\,\phi \land \Box\,\phi') \to \Box\,(\phi \land \phi')$$

Proof.

1. $\Box(p \to q) \to (\Box p \to \Box q)$ — Axiom **K**
2. $\Box(B \to A \wedge B) \to$ $(\Box B \to \Box(A \wedge B))$ — Subst. on 1.: B/p, $A \wedge B/q$
3. $A \to (B \to (A \wedge B))$ — Tautology
4. $\Box(A \to (B \to (A \wedge B)))$ — Necessitation on 3.
5. $\Box(A \to (B \to (A \wedge B))) \to$ $(\Box A \to \Box(B \to (A \wedge B)))$ — Subst. on 1.
6. $\Box A \to (\Box B \to (A \wedge B)))$ — MP on 4. and 5.
7. $\Box A \to (\Box B \to \Box(A \wedge B))$ — Prop. Logic on 6. and 2.
8. $(\Box A \wedge \Box B) \to \Box(A \wedge B)$ — Prop. Logic on 7.

Here we have used two **meta steps**, namely 7. and 8. Both can be replaced by proofs (several steps) to obtain a proof wrt. Definition 3.13. □

Theorem 3.15 (A Proof from $\{K\}$)

$\vdash_{\{K\}} \Box(\varphi_1 \wedge \ldots \wedge \varphi_n) \leftrightarrow (\Box\varphi_1 \wedge \ldots \wedge \Box\varphi_n)$.

Proof.

Using Lemmas on Slide 194 and Lemma 3.14, we have the equivalence of the formulae for $n = 2$. Then we use this equivalence $n - 1$ times and get the result. $\Box$

Finally, a modal logic is a *set of formulae* with the following properties

Definition 3.16 (Modal Logics)

A modal logic Λ is a set of modal formulas that

1. contains all propositional tautologies,
2. is closed under modus ponens, and
3. is closed under uniform substitution.

Elements of $\varphi \in \Lambda$ are called theorems of Λ, we write $\vdash_\Lambda \varphi$.

How many formulae does a modal logic contain?

Can a modal logic be represented by a *finite* set?

The following result is obvious: A set of formulae Λ is a **modal logic** iff Λ is **closed under all Λ-provable** formulae.

Definition 3.17 (Axiomatization)

Let Λ be a modal logic and Σ be a proof system over Σ such that Λ and the set of all Σ-provable formulae coincide. Then, we say that Σ is an *axiomatization of Λ*. If Σ is finite we call it a *finite axiomatization*.

Note that we do not require an axiomatization for the underlying propositional tautologies: We simply took all of them (in part 2 of Definition 3.13).

Provable formula

Note that if a formula can be proved in a calculus Σ, then its proof requires only finitely many formulae. Even if we consider infinitely man axioms, a proof can only require finitely many of them.

Sound & Completeness

Definition 3.18 (Semantic Set $\Lambda_{\mathcal{F}}$)

Let $\mathcal{F}$ be a class of **frames** (or **models**). We define $\Lambda_{\mathcal{F}}$ to be the set of all valid formulae over $\mathcal{F}$, i.e.

$$\Lambda_{\mathcal{F}} = \{\varphi \mid \mathfrak{F} \models \varphi \text{ for all } \mathfrak{F} \in \mathcal{F}\}$$

Is $\Lambda_{\mathcal{F}}$ a modal logic?

In the following we are interested in the following question:

Are there syntactic mechanisms for generating $\Lambda_{\mathcal{F}}$?

($\mathcal{F}$ is a class of frames)

Definition 3.19 (Soundness)

Let $\mathcal{F}$ be a class of frames or models. A modal logic Λ is sound with respect to $\mathcal{F}$ if for all formulae φ

$$\vdash_\Lambda \varphi \text{ implies } \models_\mathcal{F} \varphi.$$

Recall that $\models_\mathcal{F} \varphi$ means for all $\mathfrak{F} \in \mathcal{F}: \ \mathfrak{F} \models \varphi$.
Recall also that in case $\mathfrak{F}$ is a frame, $\mathfrak{F} \models \varphi$ means that φ is true in all models that can be built on $\mathfrak{F}$.

Exercise: Show that this definition is equivalent to the following: Λ is sound wrt. $\mathcal{F}$ iff $\Lambda \subseteq \Lambda_\mathcal{F}$.

Definition 3.20 (Completeness)

Let $\mathcal{F}$ be a class of frames or models.
A logic Λ is **strongly complete** with respect to $\mathcal{F}$ if for any set of formulae $\Gamma \cup \{\varphi\}$ it holds that

$$\Gamma \models_{\mathcal{F}} \varphi \text{ implies } \Gamma \vdash_{\Lambda} \varphi.$$

A logic Λ is **weakly complete** with respect to $\mathcal{F}$ if for any formula φ,

$$\models_{\mathcal{F}} \varphi \text{ implies } \vdash_{\Lambda} \varphi.$$

- All the logics that we consider in this lecture are strongly complete wrt. a certain class of frames. And thus they are also weakly complete.
- But there are also logics that are only weakly complete wrt a class of frames and not strongly complete (see Example 3.56).

Exercise: Show that weak completeness

1. is a special case of strong completeness.
2. is equivalent to the following:
   Λ is weakly complete wrt. $\mathcal{F}$ iff $\Lambda_{\mathcal{F}} \subseteq \Lambda$.

3.2 Normal Modal Logics

Normal Modal Logics

- We have defined **modal logics** in Definition 3.16.
- We noticed that **Axiom K**: $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ **is true in all Kripke frames**.
- We also showed that **necessitation is a valid rule for Kripke frames**.

Definition 3.21 (Normal Modal Logic)

A **modal logic** Λ is called *normal* if it contains the formula

$$(\mathbf{K}) \quad \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$$

and is closed under *necessitation*.

This is how we defined a *modal proof* in Definition 3.13: With Σ containing $\mathbf{K}$ and all three rules of inference.

Lemma 3.22 ($\mathbb{K}$)

There is a smallest normal modal logic*. We call it* $\mathbb{K}$*, in honor of Saul Kripke.*

Proof.

The intersection of two normal modal logics still satisfies all the properties of a normal modal logic (closure properties). □

Is there also a greatest normal modal logic?

Definition 3.23 (Consistency, Λ-consistent)

A modal logic Λ is **consistent**, if it does not contain a formula ϕ and its negation $\neg\phi$.
A set of formulae Φ is Λ-**consistent**, if the modal logic defined by Λ extended by the axioms Φ is consistent.

We call a logic that is not consistent also **inconsistent**.

The following is easy to show purely syntactically.

Lemma 3.24 (Ex Falso Quodlibet)

The following is equivalent:

- Λ *is inconsistent.*
- Λ *is the set of all formulae.*

Lemma 3.25

For any normal modal logic Λ *the following implication holds:*

$$\vdash_\Lambda \varphi \leftrightarrow \psi \textit{ implies } \vdash_\Lambda \Box\varphi \leftrightarrow \Box\psi.$$

Corollary 3.26

For any normal modal logic Λ *the following implication holds:*

$$\vdash_\Lambda \varphi \leftrightarrow \psi \textit{ implies } \vdash_\Lambda \Diamond\varphi \leftrightarrow \Diamond\psi.$$

Proof.

Wlog it suffices to show $\vdash_\Lambda \Box\,\varphi \rightarrow \Box\,\psi$ (by symmetry). From the assumption it follows that $\vdash_\Lambda \varphi \rightarrow \psi$ and thus, by necessitation, $\vdash_\Lambda \Box\,(\varphi \rightarrow \psi)$. By normality, we have $\vdash_\Lambda \Box\,\varphi \rightarrow \Box\,\psi$.
The corollary follows by purely propositional logic transformations ($\neg\varphi$ instead of φ and $\neg\psi$ instead of ψ) and the fact that $\varphi \leftrightarrow \psi$ is equivalent to $\neg\varphi \leftrightarrow \neg\psi$) and the definition of $\Diamond$. $\Box$

In fact, the proof of the last lemma shows the following.

Lemma 3.27

For any normal modal logic Λ*, the following holds. Let* Ψ *be a formula containing the constant* p*, let* φ, φ' *be two arbitrary sentences and let* σ_1, σ_2 *be two substitutions with* $\sigma_1(\mathsf{p}) = \varphi$, $\sigma_2(\mathsf{p}) = \varphi'$*. Then*

$$\text{If } \vdash_\Lambda \varphi \leftrightarrow \varphi', \text{ then } \vdash_\Lambda \psi\sigma_1 \leftrightarrow \psi\sigma_2.$$

Proof.

This is because each occurrence of $\Box\,\varphi$ can be equivalently replaced by $\Box\,\varphi'$ recursively in the formula. □

Formalizing in K4

In fact, the proof of the last lemma can even be formalized within the modal logic K4.

Theorem 3.28 (Formalization Within K4)

Let Ψ *be a formula containing the constant* p*, let* ϕ, ϕ' *be two arbitrary sentences and let* σ_1, σ_2 *be two substitutions with* $\sigma_1(\mathsf{p}) = \varphi$, $\sigma_2(\mathsf{p}) = \varphi'$. *Then:*

$$\vdash_{\{K4\}} \Box\,(\varphi \leftrightarrow \varphi') \;\rightarrow\; \Box\,(\psi\sigma_1 \leftrightarrow \psi\sigma_2)$$

This shows that modal logics can be used to formalize **statements about modal logics**. This is in particular interesting, when we formalize statements about provability of certain statements in Peano arithmetic.

How does $\mathbf{K}$ look like?

Are there other sensible logics?

Theorem 3.29 (Completeness of System K)

System K *is sound and strongly complete with respect to* arbitrary Kripke frames.

The proof follows later.

How can we use this important result to show that the formula $\Box\, p \rightarrow p$ can not be proved from **K**?

We have to construct a countermodel: take two worlds w_1, w_2 such that $w_1 \mathcal{R} w_2$ and in w_1 p is false, but in w_2 p is true.

We can define various normal modal logics by appropriate axioms.

Definition 3.30 (K, KT, B, K4, S4, S5, G)

We define the following normal modal logics by stating their axioms:

K: $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$.

KT: **K** and **T**: $\Box p \rightarrow p$.

B=KTB: **K**, **T** and **B**: $p \rightarrow \Box \Diamond p$.

K4: **K** and **4**: $\Box p \rightarrow \Box \Box p$.

S4=KT4: **K**, **T**, and **4**.

S5=KT5: **K**, **T**, and **5**: $\Diamond p \rightarrow \Box \Diamond p$.

G=KL: **K** and **L**: $\Box(\Box p \rightarrow p) \rightarrow \Box p$.

B stands for **Brouwer**, the famous dutch mathematician.

Characterisation of G

What are appropriate frames for **G** , i.e. analogous properties to Lemma 3.2?

It is easy to see that

$(W, \mathcal{R}) \models \Box(\Box \mathrm{p} \rightarrow \mathrm{p}) \rightarrow \Box \mathrm{p}$ iff $\forall S \subseteq W \forall w \in W$

$$\forall x(w\mathcal{R}x \rightarrow (\forall y(x\mathcal{R}y \rightarrow y \in S) \rightarrow x \in S)) \rightarrow \forall x(w\mathcal{R}x \rightarrow x \in S).$$

Lemma 3.31 (Appropriate Frames for G)

$(W, \mathcal{R}) \models \Box(\Box \mathrm{p} \rightarrow \mathrm{p}) \rightarrow \Box \mathrm{p}$ *iff* $\mathcal{R}$ *is* transitive *and there is no infinite sequence* $w_0, w_1, \ldots$ *such that* $w_0\mathcal{R}w_1,\ w_1\mathcal{R}w_2, \ldots$.

Proof (of Lemma 3.31) (1).

$\Rightarrow$:

1. We show that $\mathcal{R}$ is transitive. Let $S_w := \{a : a\mathcal{R}w\} \cap \{a : \forall b(a\mathcal{R}b \rightarrow w\mathcal{R}b\}$ for $w \in W$. It suffices to show that for every a such that $w\mathcal{R}a$ and $a \in S_w$ and for every b with $a\mathcal{R}b$ we have $w\mathcal{R}b$.
2. The second condition. Suppose there is such an infinite sequence and consider the set S' consisting of these elements. Let $w \in S'$. Then there is a $z \in S'$ with $w\mathcal{R}z$ and we consider $S = W \setminus S'$. For all x, if $w\mathcal{R}x$ and $\forall y(x\mathcal{R}y \rightarrow y \in S)$ then $x \in S$ (why?). But this is exactly the formula $\Box(\Box \mathsf{p} \rightarrow \mathsf{p})$. Therefore (here we use the formula $\Box(\Box \mathsf{p} \rightarrow \mathsf{p}) \rightarrow \Box \mathsf{p})$, we also have $\forall x(w\mathcal{R}x \rightarrow x \in S)$, but this is a contradiction because $z \in S'$.

□

Proof (of Lemma 3.31) (2).

$\Leftarrow$: Assume the formula is not true. I.e. there is $w, x' \in W$, $w\mathcal{R}x'$ and $x' \models \neg$p and $w \models \Box(\Box\,\text{p} \rightarrow \text{p})$. Thus $x' \models \Box\,\text{p} \rightarrow \text{p}$, thus $x' \models \Diamond\neg$p. So there is a x'' with $x'\mathcal{R}x''$ and $x'' \models \neg$p. Because of transitivity, $w\mathcal{R}x''$ and thus $x'' \models \Box\,\text{p} \rightarrow \text{p}$ and we get an infinite sequence x_i with $x_i\mathcal{R}x_{i+1}$ which is a contradiction. □

Theorem 3.32 (Subsystems of KDT45B)

Let $\mathbf{X}$ *be any subset of* $\{\mathbf{D}, \mathbf{T}, \mathbf{4}, \mathbf{5}, \mathbf{B}\}$ *and let* $\mathcal{X}$ *be any subset of* $\{$*serial*, *reflexive*, *transitive*, *euclidean*, *symmetric*$\}$ **corresponding to** $\mathbf{X}$.
Then $\mathbf{K} \cup \mathbf{X}$ *is* sound and strongly complete *with respect to Kripke frames the accessibility relation of which satisfies* $\mathcal{X}$.

Note, that this theorem also holds for any modal similarity type τ and the appropriate Kripke models (frames). The properties above are then formulated for all operators in τ. This will be important in the next sections when we talk about logics with several agents.

Proof.

$\rightsquigarrow$ Will be done later on Slide 260. □

Corollary 3.33 (Completeness of KT45 (S5))

*System **S5=KT5** is sound and strongly complete with respect to Kripke frames $(W, \mathcal{R})$ where $\mathcal{R}$ is an equivalence relation.*

Corollary 3.34 (Completeness of KT4 (S4))

*System **S4=KT4** is sound and strongly complete with respect to Kripke frames $(W, \mathcal{R})$ where $\mathcal{R}$ is a preorder.*

The Case of System G

Theorem 3.35 (Completeness of G)

System **G** *is sound and* weakly complete *with respect to Kripke frames that are transitive and* well-capped*: there is no infinite sequence* $w_0, w_1, \ldots$ *such that* $w_0 \mathcal{R} w_1,\ w_1 \mathcal{R} w_2, \ldots$*.*

Proof.

$\leadsto$ Will be done later on Slide 265. □

Note: GT is inconsistent.

The system **GT** consisting of axioms K, T and G is inconsistent.

Apply necessitation to **T**, then apply **G** and again **T**: one can derive p for all p!

S4Grz

The following axiom plays a role in the next theorem

$$\textbf{Grz}: \ \Box(\Box(p \to \Box p) \to p) \to p$$

The proof of the following theorem will be given in the next section.

Theorem 3.36 (Completeness of S4Grz)

*System **S4Grz=KT4Grz** is sound and weakly complete with respect to Kripke frames $(W, \mathcal{R})$ where $\mathcal{R}$ is a finite partial order.*

Soundness of S4Grz

1. We only have to show that the axiom **Grz** holds in any finite partial ordering (why?).
2. We assume a transitive, reflexive relation on a finite set W of worlds, and a $w \in W$ such that **Grz** does not hold in $\mathfrak{M}, w$. **Then we show that $\mathcal{R}$ is not antisymmetric.**
3. We define a sequence $w_0, w_1, \ldots, w_i, \ldots$ of worlds s.t. for all i: $w_i \neq w_{i+1}$ and $w_i \mathcal{R} w_{i+1}$.

Then we are done: As W is finite, there must be j, i with $w_i = w_j$, $i \neq j$ and $w_i \mathcal{R} w_j$. Thus $\mathcal{R}$ is not antisymmetric.

Construction of the Sequence

1. $w_0 := w$. Thus p does not hold in w_0, but $\Box(\Box(p \to \Box p) \to p)$ does.
2. We want to make sure that this is true in all *even* worlds w_{2i} and do this by induction.
3. By reflexivity, in these even worlds also holds $\Box(p \to \Box p) \to p$ and thus $\Box(p \to \Box p)$ does not hold.
4. So for some w_{2i+1} with $w_{2i}\mathcal{R}w_{2i+1}$, $w_{2i+1} \not\models p \to \Box p$, thus $w_{2i+1} \models p$ and $w_{2i+1} \not\models \Box p$ and $w_{2i+1} \neq w_{2i}$.
5. As $w_{2i+1} \not\models \Box p$, there is some w_{2i+2} with $w_{2i+1}\mathcal{R}w_{2i+2}$ and $w_{2i+2} \not\models p$, thus $w_{2i+1} \neq w_{2i+2}$.
6. By transitivity, the formula $\Box(\Box(p \to \Box p) \to p)$ holds in w_{2i+2} and this completes the construction of the sequence.

Theorem 3.37 (Some Normal Modal Logics)

We have the following strict relationships *between the introduced logics.*

$$
\begin{array}{ccccc}
\mathbf{K} & \subsetneq & \mathbf{K4} & \subsetneq & \mathbf{G} \\
\rotatebox{-90}{$\subsetneq$} & & \rotatebox{-90}{$\subsetneq$} & & \\
\mathbf{KT} & \subsetneq & \mathbf{S4} & \subsetneq & \mathbf{S4Grz} \\
\rotatebox{-90}{$\subsetneq$} & & \rotatebox{-90}{$\subsetneq$} & & \rotatebox{-90}{$\subsetneq$} \\
\mathbf{B} & \subsetneq & \mathbf{S5} & \subsetneq & \mathbf{S5Grz}
\end{array}
$$

Proof.

$\rightsquigarrow$ $\rightsquigarrow$ Blackboard Most relationships are obvious. The following require some thinking: (1) **S4** $\subsetneq$ **S5**, (2) **K4** $\subsetneq$ **G**, (3) **S4** $\subsetneq$ **S4Grz**, (4) **S5** $\subsetneq$ **S5Grz**, (5) **B** $\subsetneq$ **S5**, (6) **S4Grz** $\subsetneq$ **S5Grz**: Find appropriate models. □

Lemma 3.38 (Some Properties of S5)

1. *The axiom* ***D*** *follows from* ***KT5***: $\vdash_{\mathbf{KT5}} \mathbf{D}$.
2. *The formula* $\Diamond p$ *does not follow from* ***KT5***.
3. ***KD5*** *is* **not equivalent** *to* ***K5***. *Axiom* ***D*** *does* **not follow** *from* ***K5***: $\nvdash_{\mathbf{K5}} \mathbf{D}$.

$\leadsto$ Exercise

Lewis' System S4

- In Lewis system, the modality can be interpreted as is provable.
- **S4** has the **finite model property**: If a formula ϕ is not derivable, then there is a finite countermodel.
- It is therefore decidable (why?).
- The formula $\Box(\Box(p \rightarrow \Box p) \rightarrow \Box p) \rightarrow (\Diamond\Box p \rightarrow \Box p)$ (from Dummett) is **not derivable** in **S4**.
- Consider **S4** and the McKinsey axiom $\Box\Diamond p \rightarrow \Diamond\Box p$. Then the binary relation in the corresponding frames satisfies: $\forall x \exists y (xRy \wedge \forall z (yRz \rightarrow y = z))$.

Density and Axiom 4

We consider the set of all frames $(W, \mathcal{R})$ where $\mathcal{R}$ is a **linear** ordering.

In which of these frames is $\Box$ p $\rightarrow$ $\Box$ $\Box$ p true?

In all frames where $\mathcal{R}$ is a dense linear ordering:

$$\forall x \forall y \exists z : (x\mathcal{R}y \wedge x \neq y) \rightarrow (x\mathcal{R}z \wedge z\mathcal{R}y \wedge x \neq z \wedge z \neq y)$$

$\rightsquigarrow$ $\rightsquigarrow$ Blackboard

Idea: Assume a linear, non-dense order, i.e. between w and w' ($w\mathcal{R}w'$) there are no other worlds. Then construct an appropriate Kripke model where the formula is false: Make p true **only in** w' and nowhere else. Then the formula is false in world w in the Kripke model.

3.3 Sound- and Completeness

Proofs of the Theorems

- In this section we introduce the machinery to prove our completeness theorems for normal modal logics (developed by Scott and Makinson in the 60ies).
- It is based on Definition 3.43 and the general Theorem 3.47.
- We only have to check whether the canonical model satisfies the required properties.
- This method gives completeness proofs for systems **K**, **K4**, **KT**, **S4**, **B**, **S5**.
- The systems **G** and **S4Grz** need separate treatment. We prove slightly stronger completeness theorems. This method leads to the filtration approach.

Remember Definition 3.20. Here are two simple characterizations that we use in the following.

Lemma 3.39 (Strongly, Weakly Complete)

A logic Λ is **strongly complete** *with respect to a class of structures $\mathcal{F}$ if and only if, every Λ-consistent* **set of formulae** *is satisfiable on some $\mathfrak{F} \in \mathcal{F}$.*
Λ is **weakly complete** *with respect to a class of structures $\mathcal{F}$ if and only if, every Λ-consistent* **formula** *is satisfiable on some $\mathfrak{F} \in \mathcal{F}$.*

We only show the proof for strong completeness (weak completeness follows).

Proof.

"$\Leftarrow$": Suppose Λ is not strongly complete wrt $\mathcal{F}$. Thus, there is $\Gamma \cup \{\varphi\}$ such that $\Gamma \models_{\mathcal{F}} \varphi$ and $\Gamma \not\vdash_{\Lambda} \varphi$. Then, $\Gamma \cup \{\neg\varphi\}$ is Λ-consistent but not satisfiable on $\mathcal{F}$.

"$\Rightarrow$": Let Φ is a Λ-consistent **set of formulae**. Suppose Φ is not satisfiable at some $\mathfrak{F} \in \mathcal{F}$. Then $\Phi \models_{\mathcal{F}} \mathrm{p} \wedge \neg\mathrm{p}$ for all p. But then strong completeness implies $\Phi \vdash_{\Lambda} \mathrm{p} \wedge \neg\mathrm{p}$, which is a contradiction to the consistency of Φ. □

Canonical Models

Definition 3.40 (Λ-MCS's)

A set Γ of formulae is maximal Λ-consistent (short: Λ-MCS) if Γ is Λ-consistent and any set of formulae properly containing Γ is Λ-inconsistent.

Proposition 3.41 (Properties of MCS's)

Let Λ be a logic and Γ a Λ-MCS. Then:

1. Γ *is closed under modus ponens.*
2. $\Lambda \subseteq \Gamma$.
3. Γ *is a complete theory.*
4. *For all formulae* ϕ, ψ: $\phi \vee \psi \in \Gamma$ *iff* $\phi \in \Gamma$ *or* $\psi \in \Gamma$.

Proof.

1. If not, it were not maximal.
2. If not, it were not maximal.
3. If not, it were not maximal.
4. If not, it were not maximal.

□

Theorem 3.42 (Lindenbaum's Lemma)

If Γ is a Λ-consistent set of formulae then **there exists a Λ-MCS Γ^+ such that $\Gamma \subseteq \Gamma^+$.**

Proof.

We enumerate all formulae $\phi_1, \ldots, \phi_i, \ldots$ and define $\Sigma_0 = \Sigma$,

$$\Sigma_{n+1} := \begin{cases} \Sigma_n \cup \{\Phi_n\}, & \text{if this theory is } \Lambda\text{-consistent;} \\ \Sigma_n \cup \{\neg\Phi_n\}, & \text{else.} \end{cases}$$

and $\Sigma^+ := \bigcup_{n \geq 0} \Sigma_n$. Rest on Blackboard. □

Definition 3.43 (Canonical Model)

The canonical model $\mathfrak{M}^\Lambda = (W^\Lambda, \mathcal{R}^\Lambda, V^\Lambda)$ for a modal logic Λ is defined as follows:

1. W^Λ is the set of all Λ-MCS's;
2. $\mathcal{R}^\Lambda \subseteq W^\Lambda \times W^\Lambda$ such that $\mathcal{R}^\Lambda ww'$ if $\varphi \in w'$ implies $\Diamond \varphi \in w$ for all φ. $\mathcal{R}^\Lambda$ is called canonical relation;
3. $V^\Lambda(p) := \{w \in W^\Lambda \mid p \in w\}$. V^Λ is the canonical relation.

$\mathfrak{F}^\Lambda = (W^\Lambda, \mathcal{R}^\Lambda)$ is the canonical frame of Λ.

Lemma 3.44

For any normal logic Λ:

$$\mathcal{R}^{\Lambda}ww' \text{ if, and only if } \Box\varphi \in w \text{ implies } \varphi \in w' \text{ for all } \varphi \in \Lambda.$$

Proof.

$\Rightarrow$: Let $\mathcal{R}^{\Lambda}ww'$ and $\Box\,\varphi \in w$ and $\varphi \notin w'$. We have to show a contradiction. Then $\neg\varphi \in w'$. But also $\Diamond\,\neg\varphi \in w$ (Definition of $\mathcal{R}$). Thus $\neg\Diamond\,\neg\varphi \notin w$, which means $\Box\,\varphi \notin w$, which is a contradiction.

$\Leftarrow$: Let the implication be true. Assume not $\mathcal{R}^{\Lambda}ww'$, i.e. there is ψ, $\psi \in w'$ and $\Diamond\,\psi \notin w$. Then $\Box\,\neg\psi \in w$. But then (using the very implication): $\neg\psi \in w'$, in contradiction to the consistency of w'.

$\Box$

Lemma 3.45 (Existence Lemma)

For any normal logic Λ and any state $w \in W^\Lambda$:

$\Diamond\varphi \in w$ *implies* *there is a state* $w' \in W^\Lambda$ *such that* $\mathcal{R}^\Lambda ww'$ *and* $\varphi \in w'$

Proof.

Assume $\Diamond \varphi \in w$. How to construct w'?
Let w' be **any Λ-MCS extending the set** $\{\varphi\} \cup \{\psi : \Box \psi \in w\}$. Obviously $\mathcal{R}^{\Lambda} w w'$. It remains to show that the set

$$\{\varphi\} \cup \{\psi : \Box \psi \in w\}$$

is consistent.
Assume not: then $\vdash_{\Lambda} \psi_1 \wedge \ldots \wedge \psi_n \rightarrow \neg\varphi$. But then $\vdash_{\Lambda} \Box (\psi_1 \wedge \ldots \wedge \psi_n) \rightarrow \Box \neg\varphi$, and therefore (normal ML)

$$\vdash_{\Lambda} \Box \psi_1 \wedge \ldots \wedge \Box \psi_n \rightarrow \Box \neg\varphi.$$

Thus (w is a MCS) $\Box \psi_1 \wedge \ldots \wedge \Box \psi_n \in w$, and therefore $\Box \neg\varphi \in w$, which contradicts $\Diamond \varphi \in w$. $\Box$

Lemma 3.46 (Truth Lemma)

For any normal logic Λ and any formula φ:

$$\mathfrak{M}^{\Lambda}, w \models \varphi \quad \textit{iff} \quad \varphi \in w$$

Proof.

By induction on the formula φ. Base cases and boolean combinations are trivial (see Proposition 3.41). We consider the case $\Diamond \varphi$. This is true in $\mathfrak{M}^{\Lambda}, w$ iff there is a v with $\mathcal{R}^{\Lambda}wv$ and $\mathfrak{M}^{\Lambda}, v \models \varphi$ iff there is a v with $\mathcal{R}^{\Lambda}wv$ and $\varphi \in v$ (Induction Hypothesis) only if $\Diamond \varphi \in w$.
For right to left we apply Lemma 3.45. □

Theorem 3.47 (Canonical Model Theorem)

Any modal logic is strongly complete with respect to its canonical model.

Proof.

Let Σ be a consistent set of the logic Λ. Extend it to a MCS Σ^+ (Theorem 3.42). But then (using Lemma 3.46)

$$\mathfrak{M}^\Lambda, \Sigma^+ \models \Sigma.$$

□

Completeness for Systems K, . . . , S5

Theorem 3.47 does not seem to be too powerful: It only speaks about one single model, whereas we want to prove completeness with respect to a class of frames, not a single model.

Lemma 3.48 (Strong Completeness wrt $\mathcal{F}$)

Consider any subset ***X*** *of* $\{K, D, T, B, 4, 5\}$ *and let us denote the corresponding normal modal logic by* ***X****. We also denote by* $\mathcal{F}_{\mathbf{X}}$ *the corresponding class of frames (see Lemma 3.2).*
Assume that for any ***X****-consistent set of formulae* Γ, **the canonical frame** $(W^\Gamma, \mathcal{R}^\Gamma)$ **is contained in** $\mathcal{F}_{\mathbf{X}}$.
Then ***X*** *is* strongly complete *wrt. the class of frames* $\mathcal{F}_{\mathbf{X}}$.

Proof.

According to Lemma 3.39, we have to find a model for Γ. We just take the canonical model $\mathcal{M}$ and any **X**-MCS extending Γ: Γ^+. By Theorem 3.47, this model satisfies Γ. □

We now prove Theorem 3.32 by applying Lemma 3.48. We have to show for each of our axioms that the canonical model has the appropriate property (as given in Lemma 3.2).

K: There is nothing to prove. We have already shown in Lemma 3.2 that this axiom holds in all frames.

T: Obviously, $\mathcal{R}^{\mathsf{X}}ww$ because **X** contains $\Box\, \mathsf{p} \rightarrow \mathsf{p}$.

4: Assume $\mathcal{R}^{\mathsf{X}}wx$ and $\mathcal{R}^{\mathsf{X}}xy$. We have to show that $\mathcal{R}^{\mathsf{X}}wy$. **Suppose** $\Box\, \mathsf{p} \in w$. As $\Box\, \mathsf{p} \rightarrow \Box\,\Box\, \mathsf{p}$ is an axiom, $\Box\,\Box\, \mathsf{p} \in w$. Then $\Box\, \mathsf{p} \in x$ and therefore $\mathsf{p} \in y$.

B: Suppose $\mathcal{R}^{\mathsf{X}}wx$. We have to show that $\mathcal{R}^{\mathsf{X}}xw$, i.e. that $\Box\, \mathsf{p} \in x$ implies $\mathsf{p} \in x$. Suppose $\mathsf{p} \notin w$. Then $\neg \mathsf{p} \in w$. Using **B**, then $\Box \Diamond \neg \mathsf{p} \in w$. Thus $\Diamond \neg \mathsf{p} \in x$, and $\neg \Box\, \mathsf{p} \in x$. Therefore $\Box\, \mathsf{p} \notin x$.

5: Suppose $\mathcal{R}^{\mathsf{X}}wx$ and $\mathcal{R}^{\mathsf{X}}wy$. We have to show that if $\Box\, \mathsf{p} \in x$, then $\mathsf{p} \in y$. Assume $\mathsf{p} \notin y$. Then $\Box\, \mathsf{p} \notin w$, $\neg \Box\, \mathsf{p} \in w$, and thus $\Diamond \neg \mathsf{p} \in w$. Using the axiom, we conclude $\Box \Diamond \neg \mathsf{p} \in w$, $\Diamond \neg \mathsf{p} \in x$. Thus $\neg \Box\, \mathsf{p} \in x$, and therefore $\Box\, \mathsf{p} \notin x$.

$\leadsto$ Rest on Blackboard, note that also **any combinations** of these axioms go through without any problems.

- The previous proofs show that: the canonical frame of each normal logic containing axiom T is reflexive.
- And appropriate statements for all other axioms, as well as combinations of them.
- Unfortunately, this does not work for axioms **G** or **Grz**. The canonical model of **G** is not well capped. The relation in the canonical model of **Grz** is not a finite partial order.
- But we can still use the canonical models: **We just need to transform them**!

Completeness for System G

We prove the following theorem, that is **stronger** than Theorem 3.35

Theorem 3.49 (Completeness of G, 2. Version)

System **G** *is sound and complete with respect to Kripke frames that are* finite, transitive, *and* well-capped: *There is no infinite sequence* $w_0, w_1, \ldots$ *such that* $w_0 \mathcal{R} w_1,\ w_1 \mathcal{R} w_2, \ldots$*.*

Note that a relation is called irreflexive, when there is no w with $w\mathcal{R}w$. The existence of such a w would immediately result in an infinite sequence falsifying the well-capped property.

Lemma 3.50 (Finite Transitive Relations)

Let $\mathcal{R}$ be a binary, transitive relation on a finite set. Then the following are equivalent:

- *$\mathcal{R}$ is irreflexive,*
- *$\mathcal{R}$ is well-capped.*

Theorem 3.51 (Completeness of G, 3. Version)

*System **G** is sound and complete with respect to Kripke frames that are finite, transitive, and irreflexive.*

Soundness follows trivially from Lemma 3.31 on Slide 234.

Are the two theorems also equivalent when we remove **finite**?

Completeness for System G: $\mathcal{M}_1$

We prove Theorem 3.51 through a series of lemmas and models $\mathcal{M}_i$.
We start with the canonical model $\mathfrak{M}^{\mathsf{G}} = (W^{\mathsf{G}}, \mathcal{R}^{\mathsf{G}}, V^{\mathsf{G}})$.
Suppose that φ is not a theorem of **G**. Then there is a world t s.t. $\mathfrak{M}^{\mathsf{G}}, t \not\models \varphi$.

Ultimately, we have to find a model $\mathcal{M}_3$ which is finite, transitive and irreflexive.

$\mathfrak{M}_1$: We define $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ with: $W_1 := \{t\} \cup \{x : \mathcal{R}^{\mathsf{G}}tx\}$, and $\mathcal{R}_1, V_1$ are just the restrictions of $\mathcal{R}^{\mathsf{G}}, V^{\mathsf{G}}$ to W_1.

Unfortunately, $\mathfrak{M}_1$ is in general not finite.

Completeness for System G: $\mathfrak{M}_2$

We are now making $\mathfrak{M}_1$ finite, by identifying many worlds.
We define an equivalence relation on $\mathfrak{M}_1$ as follows: w and x are equivalent iff for every subformula φ' of φ: $\varphi' \in w$ iff $\varphi' \in x$. We denote the equivalence class of w by $\overline{w}$.
We define $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ with: $W_2 = \{\overline{w} : w \in W_1\}$,

$$\mathcal{R}_2 = \{\langle \overline{w}, \overline{x} \rangle : \text{ for every subformula } \Box\,\psi \text{ of } \varphi, \text{ if } \Box\,\psi \in w \text{ then } \Box\,\psi \in x \text{ and } \psi \in x\}$$

Finally, p $\in V_2(\overline{w})$ iff p $\in V_1(w)$ if p is a subformula of φ, and let p $\notin V_2(\overline{w})$ otherwise.
Clearly, $\mathfrak{M}_2$ is transitive and finite. And if $\mathcal{R}_1 wx$ then $\mathcal{R}_2 \overline{w}\,\overline{x}$.

Completeness for System G: $\mathfrak{M}_3$

Lemma 3.52

For all combinations ψ of subformulae from φ and each $w \in W_1$:

$$\mathfrak{M}_1, w \models \psi \ \textit{iff}\ \mathfrak{M}_2, \overline{w} \models \psi$$

Proof is by induction on ϕ.

$\mathfrak{M}_3$: The problem with $\mathfrak{M}_2$ is that it is not irreflexive. We will define a subrelation $\mathcal{R}_3$ that is irreflexive. Our model is then defined as

$$\mathfrak{M}_3 = (W_2, \mathcal{R}_3, V_2).$$

Completeness for System G: $\mathfrak{M}_3$

$\mathcal{R}_3$: In order to define $\mathcal{R}_3$, we need the following:

$$\overline{w} \sim \overline{x} \text{ if } \begin{cases} \text{if } \overline{w} \sim \overline{x}\text{; or} \\ \text{if } \mathcal{R}_2\overline{wx} \text{ and } \mathcal{R}_2\overline{xw}. \end{cases}$$

$\sim$ is an equivalence relation because $\mathcal{R}_2$ is transitive.

L_S: Given $S \subseteq W$ such that for some $w \in W_1$ $S = \{\overline{y} : \overline{w} \sim \overline{y}\}$, we fix a **strict, linear ordering L_S of S**.

$\mathcal{R}_3$: $\mathcal{R}_3$ consists of all pairs $\langle \overline{w}, \overline{x} \rangle$ such that either (1) $\overline{w} \not\sim \overline{x}$ and $\mathcal{R}_2\overline{w}, \overline{x}$, or (2) $\overline{w} \sim \overline{x}$ and $\overline{w}L_S\overline{x}$, where $S = \{\overline{y} : \overline{w} \sim \overline{y}\}$.

Properties: **$\mathcal{R}_3$ is a subrelation of $\mathcal{R}_2$** because $\mathcal{R}_2$ is transitive. **$\mathcal{R}_3$ is also transitive and irreflexive.**

Putting It All Together

Lemma 3.53

For all combinations ψ *of subformulae from* φ *and each* $w \in W_1$*:*

$$\mathfrak{M}_2, w \models \psi \text{ iff } \mathfrak{M}_3, \overline{w} \models \psi$$

Proof is by induction on ϕ.

Filtrations

What we have done on the last few slides is a general method that can be applied very often: the filtration method. It requires the **equivalence of models wrt. subformulae of a given formula**. In the last section of this chapter, we investigate this approach more thoroughly.

Completeness for System S4Grz

Theorem 3.54 (Completeness of S4Grz)

*System **S4Grz=KT4Grz** is sound and complete with respect to Kripke frames $(W, \mathcal{R})$ where $\mathcal{R}$ is a finite partial order.*

The proof follows the proof of Theorem 3.49. It is not trivial.

There is a strong relation between systems **G** and **S4Grz**. We define a translation ^{t} from modal formulae to modal formulae as follows:

$$\begin{array}{lcl} {}^t\mathsf{p} & := & \mathsf{p} \\ {}^t\neg\phi & := & \neg\,{}^t\phi \\ {}^t(\phi \vee \psi) & := & {}^t\phi \vee {}^t\psi \\ {}^t\Box\,\phi & := & \Box\,{}^t\phi \wedge {}^t\phi \end{array}$$

Theorem 3.55 (Relation between S4Grz and G)

$\vdash_{\mathbf{S4Grz}} \phi$ *iff* $\vdash_{\mathbf{G}} {}^t\phi$.

Example 3.56

We consider again the axiom **L**: $\Box(\Box p \rightarrow p) \rightarrow \Box p$ and the system **G=KL**. We have seen that this defines a normal logic. We have already shown soundness and weak completeness with respect to the class of all frames, where $\mathcal{R}$ is a finite, transitive tree. Unfortunately, we could not use Theorem 3.48 on Slide-260.

This is not by accident: **there is no class of frames whatsoever, that this logic is strongly complete to.**

Lemma 3.57 (The logic G=KL)

***G** is not sound and strongly complete to any class of frames.*

Proof.

We consider the set Φ

$$\{\Diamond\, q_1\} \cup \{\Box\,(q_i \rightarrow \Diamond\, q_{i+1}) :\ 1 \leq i\}$$

We first show that this set is consistent (by showing that each finite subset is). We leave this to the reader. We also note that **G** is consistent (why?).
Now suppose there is a class of frames $\mathcal{F}$ that **G** is strongly complete to. Then $\mathcal{F}$ is nonempty. So there is a Kripke model $\mathfrak{M}$ and world w such that $\mathfrak{M}, w \models \Phi$. But that would mean that there is an infinite path, a contradiction. □

A Few Final Remarks

- While we have defined many logics and appropriate classes of frames, this is not always possible.
- There are axioms (and therefore logics) that are not complete wrt any set of frames.
- Canonical models is a powerful method, but it does not always work.
- Most of our axioms (except **Grz** and **G**) are canonical. A formula φ is canonical wrt to a property P of a class of frames $\mathcal{F}$, if (1) φ is valid in all frames satisfying P, and (2) the canonical frame of any normal modal logic containing φ has property P.
- The problem **"Is a formula φ canonical?"** is undecidable.

Canonical formulae

We consider again the example on Slide 246. Another formulation of that result is that the formula is canonical.

Lemma 3.58

The following formulae are canonical for their respective properties:

1. ***T*, *B*, *4*, *5** *are all canonical.*
2. $\Box p \rightarrow \Box\Box p$ *is canonical for density (among all linear orderings).*
3. ***L*** *is not canonical for any property.*

3.4 Finite Models via Filtration

Let Σ be a set of modal formulae and Λ be a modal logic. Consider the problem: Is Σ satisfiable in Λ?

- We can not just enumerate all Kripke models, as there are infinitely many of them (uncountably many: $2^{\aleph_0}$).
- Suppose we know that **"if it is satisfiable, then there is a finite model"**. Then the problem above is recursively enumerable (just enumerate all finite models).
- If we know in addition an upper bound on the size of the finite model, we can come up with a decision algorithm: The problem is then recursive.
- If the logic Λ is given by a recursive set of axioms (as all our normal logics in this chapter), then the problem is **recursive** (also called decidable) (why?).

Definition 3.59 (Finite Model Property)

Let $\mathcal{M}$ be a class of τ-models. We say that τ has the finite model property with respect to $\mathcal{M}$ if every τ-formula φ satisfiable in $\mathcal{M}$ is also satisfiable in a finite model in $\mathcal{M}$.

Such a property is important for the **decidability of the satisfaction problem**. But note, that in the above definition we talk about **one single** formula.

Is this property true for predicate logic? Or can you construct a formula that is only satisfiable in infinite universes?

The following results are for the basic modal language; that is, we consider only one accessibility relation.

Definition 3.60 (Subformula)

Let $\varphi \in \mathcal{L}_{BML}$. Each formula $\psi \in \mathcal{L}_{BML}$ which occurs in φ is called a subformula of φ.

Note that φ is a subformula of φ itself.
What are the subformulae of $\Box \neg \mathrm{p} \rightarrow (\neg \Diamond \mathrm{q} \wedge \mathrm{p})$?

Definition 3.61 (Closure under Subformulae)

Let $\Sigma \subseteq \mathcal{L}_{BML}$ be a set of formulae. The closure of Σ under subformulae, $Sub(\Sigma)$, is defined as the set of all subformulae of the formulae in Σ, i.e.

$$Sub(\Sigma) = \{\psi \mid \exists \varphi \in \Sigma \text{ s.t. } \psi \text{ is a subformula of } \varphi\}$$

If $\Sigma = Sub(\Sigma)$ then we say that Σ is subformula closed.

We write $Sub(\varphi)$ for $Sub(\{\varphi\})$.

Definition 3.62 (Filtration Relation)

Let $\mathfrak{M}$ be a Kripke model and $\Sigma \subseteq \mathcal{L}_{BML}$ a **subformula closed set**. The **filtration relation over $\mathfrak{M}$ and Σ** is the relation $\leftrightsquigarrow_{\Sigma,\mathfrak{M}} \subseteq W \times W$ which is defined as follows:

$$w \leftrightsquigarrow_{\Sigma,\mathfrak{M}} w' \quad \text{iff} \quad \forall \varphi \in \Sigma \, (\mathfrak{M}, w \models \varphi \text{ iff } \mathfrak{M}, w' \models \varphi)$$

$\rightsquigarrow$ Exercise: Show that $\leftrightsquigarrow_{\Sigma,\mathfrak{M}}$ is an equivalence relation.

We use $[x]_{\Sigma,\mathfrak{M}}$ to denote the equivalence class of x wrt. to $\leftrightsquigarrow_{\Sigma,\mathfrak{M}}$, i.e.

$$[x]_{\Sigma,\mathfrak{M}} := \{y \in W_{\mathfrak{M}} \mid x \leftrightsquigarrow_{\Sigma,\mathfrak{M}} y\}$$

We omit Σ and $\mathfrak{M}$ if they are clear from context.

We identify worlds in which the same formulae wrt. Σ hold.

Definition 3.63 (Filtration)

Let $\mathfrak{M}$ be a Kripke model and $\Sigma \subseteq \mathcal{L}_{BML}$ subformula closed. A **filtration of $\mathfrak{M}$ through Σ**, $\mathfrak{M}^f_{\Sigma,\mathfrak{M}} = (W^f, \mathcal{R}^f, V^f)$, is a model which satisfies:

1. $W^f := \{[w]_{\Sigma,\mathfrak{M}} \mid w \in W_{\mathfrak{M}}\}$,
2. $\mathcal{R}_{\mathfrak{M}}ww'$ implies $\mathcal{R}^f[w]_{\Sigma,\mathfrak{M}}[w']_{\Sigma,\mathfrak{M}}$,
3. $\mathcal{R}^f[w]_{\Sigma,\mathfrak{M}}[w']_{\Sigma,\mathfrak{M}}$ implies that for all $\Diamond\varphi \in \Sigma$ it holds that $\mathfrak{M}, w' \models \varphi$ implies $\mathfrak{M}, w \models \Diamond\varphi$, and
4. $V^f(\mathrm{p}) := \{[w]_{\Sigma,\mathfrak{M}} \mid \mathfrak{M}, w \models \varphi\}$ for every $\mathrm{p} \in \Sigma$.

Example 3.64

We consider the Kripke model $\mathfrak{M}$ consisting of worlds w_i, $i = 0, 1, \ldots$ over q, p, where $V(\mathsf{p}) = \{w_i : i \in \mathbb{N} \setminus \{0\}\}$ and $V(\mathsf{q}) = \{w_2\}$. The transition relation is given by $w_0 \mathcal{R} w_1$, $w_0 \mathcal{R} w_2$, $w_2 \mathcal{R} w_3$, $w_1 \mathcal{R} w_3$, and $w_i \mathcal{R} w_{i+1}$ for $i \geq 3$.
Let $\Sigma := \{\Diamond\, \mathsf{p}, \mathsf{p}\}$.
What is the filtration of $\mathfrak{M}$ through Σ?

- **A filtration has only worlds which differ concerning the truth of formulae in Σ.**
- Note that the accessibility relation is not unique, we only require 2) and 3).
- Conditions 2) and 3) are needed to prove Theorem 3.66.
- Note that condition 3) is very similar to the appropriate condition in the canonical model (Definition 3.43).
- However, we do not know yet whether filtrations exist at all.

Proposition 3.65 (Filtrations Are Finite)

Let $\mathfrak{M}$ *be a Kripke model and* $\Sigma \subseteq \mathcal{L}_{BML}$ *be a* **finite subformula closed set**. *For any filtration* $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}$ *over* $\mathfrak{M}$ *through* Σ, *the* number of worlds *in* $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}$ is bounded by $2^{|\Sigma|}$, *i.e.* $|W_{\mathfrak{M}^f_{\Sigma,\mathfrak{M}}}| \leq 2^{|\Sigma|}$.

Proof.

$\rightsquigarrow$ Exercise. □

The filtration model was constructed in such a way, that the same formulae hold.

Theorem 3.66 (Filtration Theorem)

Let $\mathfrak{M}$ *be a Kripke model,* $\Sigma \subseteq \mathcal{L}_{BML}$ *be a* **subformula closed set**, *and* $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}$ *a* filtration of $\mathfrak{M}$. *Then, it holds that*

$$\mathfrak{M}, w \models \varphi \quad \textit{iff} \quad \mathfrak{M}^f_{\Sigma,\mathfrak{M}}, [w] \models \varphi$$

for all $\varphi \in \Sigma$ *and* $w \in W_{\mathfrak{M}}$.

Proof.

Induction on the formula φ. Base cases and boolean cases are trivial (very definition of V^f and the fact that Σ is subformula closed).

"$\Rightarrow$": Let $\mathfrak{M}, w \models \Diamond \varphi$. Then there is a world w' with $w\mathcal{R}w'$ and $\mathfrak{M}, w' \models \varphi$. By the second clause of the definition of filtration, $\mathcal{R}^f[w]_{\Sigma,\mathfrak{M}}[w']_{\Sigma,\mathfrak{M}}$. By induction hypothesis, $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}, [w'] \models \varphi$ and thus $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}, [w] \models \Diamond \varphi$.

"$\Leftarrow$": For the other direction, let $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}, [w] \models \Diamond \varphi$. Then there is w' with $\mathcal{R}^f[w]_{\Sigma,\mathfrak{M}}[w']_{\Sigma,\mathfrak{M}}$ and $\mathfrak{M}^f_{\Sigma,\mathfrak{M}}, [w'] \models \varphi$. By induction hypothesis, $\mathfrak{M}, w' \models \varphi$. By the third clause, $\mathfrak{M}, w \models \Diamond \varphi$.

□

Given a model there are various filtrations of this model. This is, because the relation $\mathcal{R}^f$ can be defined in many ways. We present two of them:

1. The second item in the definition of a filtration leads to the following, **weakest possible condition**:

$$\mathcal{R}^s[w][v] \quad \text{iff} \quad \exists w' \in [w] \exists v' \in [v] \; (\mathcal{R}w'v')$$

2. The third item in the definition of a filtration leads to the following, **strongest possible condition**:

$$\mathcal{R}^l[w][v] \quad \text{iff} \quad \forall \Diamond \varphi \in \Sigma \; (\mathfrak{M}, v \models \varphi \Rightarrow \mathfrak{M}, w \models \Diamond \varphi)$$

Lemma 3.67

Let $\mathfrak{M}$ *be a Kripke model and* $\Sigma \subseteq \mathcal{L}_{BML}$ *be a* **subformula closed set**. *Then, both* $(W_{\Sigma,\mathfrak{M}}, \mathcal{R}^s, V^f)$ *and* $(W_{\Sigma,\mathfrak{M}}, \mathcal{R}^l, V^f)$ *are* **filtrations of** $\mathfrak{M}$ **through** Σ. *Moreover:*

$$\mathcal{R}^s \subseteq \mathcal{R}^f \subseteq \mathcal{R}^l$$

for any filtration $\mathfrak{M}^f_{\mathfrak{M},\Sigma}$ *of* $\mathfrak{M}$ *through* Σ.

Proof.

That $(W_{\Sigma,\mathfrak{M}}, \mathcal{R}^s, V^f)$ and $(W_{\Sigma,\mathfrak{M}}, \mathcal{R}^l, V^f)$ are the smallest resp. largest possible filtrations follows trivially from Definition 3.63.

Why are both indeed filtrations? We have to show the following:

$\mathcal{R}^s$ satisfies condition 3 of Definition 3.63. Let $\mathcal{R}^s[w][v]$, $\Diamond\varphi \in \Sigma$ and $\mathfrak{M}, v \models \varphi$ (we have to show $\mathfrak{M}, w \models \Diamond\varphi$). By definition of $\mathcal{R}^s$, there are $w_0, v_0 \in W$ with $\mathcal{R}w_0v_0$. Therefore $\mathfrak{M}, v_0 \models \varphi$ (filtration), thus $\mathfrak{M}, w_0 \models \Diamond\varphi$ (Kripke model) and thus $\mathfrak{M}, w \models \Diamond\varphi$ (filtration).

$\mathcal{R}^l$ satisfies condition 2 of Definition 3.63. This is proved along the same lines.

□

Theorem 3.68 (Finite Model Property)

Let $\varphi \in \mathcal{L}_{BML}$. If φ is satisfiable on some Kripke model, then it is also satisfiable on a finite Kripke model *which has at most 2^m worlds where $m = |Sub(\{\varphi\})|$.*

Proof.

If φ is satisfiable on some Kripke model, then we build the subformula closed set Σ of φ. Then we consider any filtration of the model through Σ. This is a **finite Kripke model** and the bound follows from Proposition 3.65. □

One important question remains: **Which properties does the relation $\mathcal{R}^f$ inherit from $\mathcal{R}$?**
For example, if $\mathcal{R}$ is transitive, is $\mathcal{R}^f$ transitive as well?

Proposition 3.69

Let $\mathfrak{M}$ be a Kripke model and $\Sigma \subseteq \mathcal{L}_{BML}$ be a **subformula closed set**. *Let $\mathcal{R}^t$ be defined as follows:*

$$\mathcal{R}^t[w][v]$$
iff
$$\forall \varphi \in \mathcal{L}_{BML} \left((\Diamond \varphi \in \Sigma \text{ and } \mathfrak{M}, v \models \varphi \vee \Diamond \varphi) \Rightarrow \mathfrak{M}, w \models \Diamond \varphi \right)$$

Then, $(W_{\Sigma,\mathfrak{M}}, \mathcal{R}^t, V^f)$ is a **filtration** *of $\mathfrak{M}$ through Σ and $\mathcal{R}^t$ is* **transitive** *if $\mathcal{R}$ was transitive.*

- The smallest possible filtration $\mathcal{R}^s$ preserves symmetry.
- However, there is **no general method** to ensure properties of $\mathcal{R}^f$.

4. Public Announcement Logic

4 Public Announcement Logic

- Epistemic Language
- Public Announcement Language
- Properties of Public Accouncements
- Epistemic Logic
- Public Announcement Logic

Content of this Chapter

In this chapter we come back to the muddy children puzzle from the first chapter. Epistemic logic is a multi modal logic allowing to talk about knowledge **within the language**. We define a logic of common knowledge S5C. Then, we introduce the public announcement language and consider some of its properties. Finally, we prove sound- and completeness results for both logics.

4.1 Epistemic Language

Epistemic Logic allows us to reason about knowledge.
It is a particular instance of our modal language. Instead of $[i]$ we write K_i.

Then, for each agent i, $K_i\varphi$ is interpreted as "agent i knows (that) φ".

Example 4.1

$\mathsf{p} \land \neg K_a\mathsf{p}$: "p is true but agent a does not know it."

$\neg K_b K_c\mathsf{p} \land \neg K_b\neg K_c\mathsf{p}$: "Agent b does not know whether agent c knows p."

Example: Alternating Bit Protocol

Example 4.2

We have two agents:

- Agent a reads a tape $X = \langle x_0, x_1, \ldots \rangle$ and sends all the inputs to agent b, and
- agent b writes down everything it receives on an output tape Y.

The **communication is not trustworthy**, i.e, there is no guarantee that all messages arrive.

Possible properties:

Fairness: If one repeats sending a certain message, it will **eventually arrive**.

Safety: At any moment, Y is a prefix of X, i.e. b will **only write a correct initial part of X into Y**.

Liveness: Every x_i will eventually be written as y_i on Y.

Assume that **fairness** holds. Can we write a protocol (or program) that satisfies the two constraints, **safety** and **liveness**?

Safety can be easily achieved by allowing b to never write anything. But then, the **liveness** property enforces that every bit of X should eventually appear in Y.

Protocol for a

$i := 0$
while true **do**
  read x_i
  repeat
    send x_i
  until $K_a K_b(x_i)$
  repeat
    send $K_a K_b(x_i)$
  until $K_a K_b K_a K_b(x_i)$
  $i := i + 1$
end while

Protocol for b

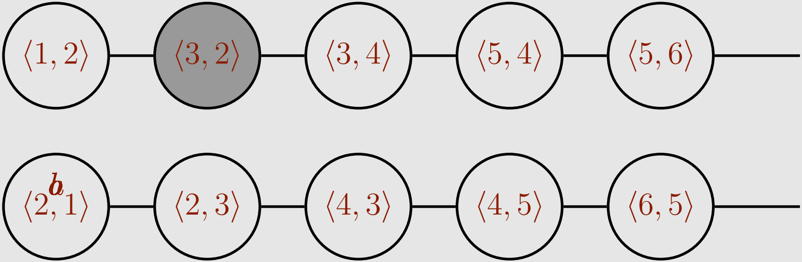
if $K_b(x_0)$ **then**
  $i := 0$
  while true **do**
    write x_i
    repeat
      send $K_b(x_i)$
    until $K_b K_a K_b(x_i)$
    repeat
      send $K_b K_a K_b(x_i)$
    until $K_b(x_{i+1})$
    $i := i + 1$
  end while
end if

Is knowledge of "depth 4" neccessary? $\rightsquigarrow$ Exercise

Example 4.3 (Consecutive Numbers)

- Assume two agents a and b. They see a number on each other's head, and those numbers are consecutive numbers n and $n+1$ for a certain $n \in \mathbb{N}$. This is common knowledge.
- a_n and b_m with $n, m \in \mathbb{N}$ means that the number on the head of a is n, on the head of b is m.
- Suppose a_3 and b_2 are true.

Let for all $w = \langle i, j \rangle$ with $i, j \in \mathbb{N}$ be $V(\mathsf{a}_i) = \{\langle i, j \rangle \mid j \in \mathbb{N}_0\}$ and $V(\mathsf{b}_j) = \{\langle i, j \rangle \mid i \in \mathbb{N}_0\}$.



$\rightsquigarrow$ Blackboard

- $\mathfrak{M}, \langle 3, 2 \rangle \models K_a \mathsf{b}_2$
- $\mathfrak{M}, \langle 3, 2 \rangle \models K_a(\mathsf{a}_1 \vee \mathsf{a}_3)$
- $\mathfrak{M}, \langle 3, 2 \rangle \models K_a K_b(\mathsf{b}_0 \vee \mathsf{b}_2 \vee \mathsf{b}_4)$
- $\mathfrak{M}, \langle 3, 2 \rangle \models K_b K_a K_b(\mathsf{a}_1 \vee \mathsf{a}_3 \vee \mathsf{a}_5)$

Common Knowledge

Common knowledge of p in a group of agents B means: All the agents in B know p, they also know that they know p, they also know that they all know that they know p, ... and so on.

So far we have no way to express common knowledge in our language, e.g. we cannot define a well-formed formula that describes that for all $n \in \mathbb{N}$ we have

$$(K_a K_b)^n \varphi$$

where $(K_a K_b)^n = K_a K_b (K_a K_b)^{n-1}$.

Example 4.4 (Byzantine Generals)

Imagine two allied generals, a and b, standing on two mountains, with their enemy in the valley between them. It is generally known that a and b together can defeat the enemy, but if only one of them attacks, he will lose the battle.

- General a sends a messenger to b with the message m about the exact time when the battle starts.
- However, it is not guaranteed that the messenger will arrive.

- Suppose, the messenger does reach the other side: K_bm and K_bK_am
- Will it be a good idea to attack? No.
- a does not know that b knows.
- b sends a messenger with an acknowledgment m′: $K_aK_bK_a$m
- Will it be a good idea to attack? No.
- b does not know whether m′ has arrived.
- We can do this forever,
- . . . common knowledge will never be established!

Group Notions of Knowledge

What about knowledge of groups?

Definition 4.5 (Epistemic Language)

Let $\mathcal{P}rop$ be a set of propositions and $\mathcal{A}g$ be a finite set of (names for) agents. The basic epistemic language with common and distributed knowledge $\mathcal{L}_{CD}(\mathcal{P}rop, \mathcal{A}g)$ consists of all formulae defined by the following grammar:

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid D_B\varphi \mid C_B\varphi$$

where $\mathsf{p} \in \mathcal{P}rop$ and $B \subseteq \mathcal{A}g$. The sublanguage without distributed knowledge is denoted by $\mathcal{L}_C(\mathcal{P}rop, \mathcal{A}g)$.

$C_B\ (\varphi)$: "Agents in B have common knowledge that φ"
$D_B\ (\varphi)$: "Agents in B have distributed knowledge that φ"

Definition 4.6 (Macros)

We define the following syntactic constructs as macros ($p \in \mathcal{P}rop$, $i \in \mathcal{A}g$):

$$
\begin{aligned}
\bot &:= p \land \neg p \\
\top &:= \neg\bot \\
\varphi \land \psi &:= \neg(\neg\varphi \lor \neg\psi) \\
\varphi \to \psi &:= \neg\varphi \lor \psi \\
\varphi \leftrightarrow \psi &:= (\varphi \to \psi) \land (\psi \to \varphi) \\
K_i\varphi &:= C_{\{i\}}\varphi \\
\hat{K}_i\varphi &:= \neg K_i\neg\varphi
\end{aligned}
$$

Note: The K_i operator is defined by the $C_{\{i\}}$ operator.

$\hat{K}_i\ (\varphi)$: "Agent i considers it possible that φ".

Definition 4.7 (Macros for Groups of Agents)

For any group $B \subseteq \mathcal{A}g$ of agents $\mathcal{A}g$ we define:

$$E_B\varphi \;:=\; \bigwedge_{b\in B} K_b\varphi := \bigwedge_{b\in B} C_{\{b\}}\varphi$$

$$\hat{E}_B\varphi \;:=\; \neg E_B\neg\varphi := \bigvee_{b\in B} \hat{K}_b\varphi := \bigvee_{b\in B} \neg C_{\{b\}}\neg\varphi$$

$E_B\ (\varphi)$: “Everybody in B knows φ”

$\hat{E}_B\ (\varphi)$: “At least one individual in B considers φ a possibility”

Semantics

The semantics for our modal language $\mathcal{L}_{CD}$ can be obtained without adding additional features to our epistemic models. We only have to define the semantics of the common knowledge operator. For this we need the transitive closure: see also Slide 201.

Definition 4.8 ((Reflexive) Transitive Closure)

Let S be a set and $\mathcal{R} \subseteq S \times S$ be a relation. The transitive closure of $\mathcal{R}$, denoted $\mathcal{R}^+$, is the smallest relation $\mathcal{R}'$ such that:

1. $\mathcal{R} \subseteq \mathcal{R}'$,
2. $\mathcal{R}'$ is transitive.

If we also require that for all x, $\mathcal{R}^+xx$ holds, we obtain the reflexive transitive closure, denoted by $\mathcal{R}^*$.

Remark 4.9

1. *If $\mathcal{R}$ is reflexive, then $\mathcal{R}^+ = \mathcal{R}^*$.*
2. *$\mathcal{R}^+ xy$ iff either $x = y$ and $\mathcal{R}xy$ or else for some $n > 1$ there is a sequence $x_1, x_2, \ldots, x_n$ such that $x_1 = x, x_n = y$ and for all $i < n$, $\mathcal{R}x_i x_{i+1}$.*

Example 4.10

- $\mathcal{R} = \{(a,b),(b,c)\}$
- $\mathcal{R}^+ = \{(a,b),(b,c),(a,c)\}$
- $\mathcal{R}^* = \{(a,b),(b,c),(a,c),(a,a),(b,b),(c,c)\}$

Definition 4.11 (Epistemic Model)

Consider a Kripke model $(W, \{\sim_i \mid i \in \mathcal{A}g\}, V)$. We call such a model epistemic model or S5-model, if all the relations $\sim_i$ are equivalence relations, i.e. $\sim_i$ is transitive, reflexive and symmetric.

$w \sim_i w'$ expresses that for agent i the two worlds w, w' are indistinguishable.

Definition 4.12

For $A \subseteq \mathcal{A}g$ we define

- $\sim_A^E$ as $\bigcup_{a \in A} \sim_a$, (everybody knows)
- $\sim_A^D$ as $\bigcap_{a \in A} \sim_a$, and (distributed knowledge)
- $\sim_A^C$ as $(\sim_A^E)^*$ (common knowledge).

Definition 4.13 (Semantics)

Let $\mathfrak{M} = (W, \{\sim_i \mid i \in \mathcal{A}g\}, V)$ be a Kripke model, $w \in W_{\mathfrak{M}}$, and $\varphi \in \mathcal{L}_C$. φ is said to be **true** or **satisfied in $\mathfrak{M}$ and world w**, written as $\mathfrak{M}, w \models \varphi$, if the semantics of $\mathcal{L}_{\mathsf{ML}}$ is extended by the following clause:

$\mathfrak{M}, w \models C_B\varphi$ iff for all worlds $w' \in W$ with $\sim_B^C ww'$ we have that $\mathfrak{M}, w' \models \varphi$,

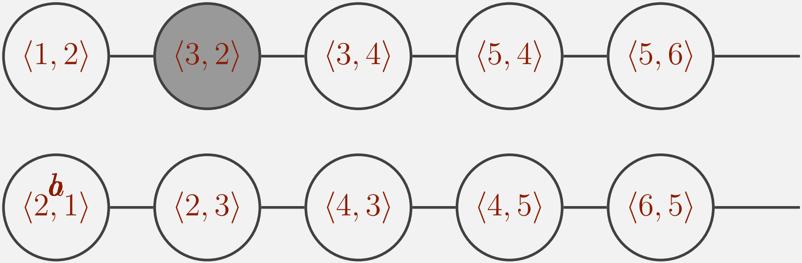
$\mathfrak{M}, w \models D_B\varphi$ iff for all worlds $w' \in W$ with $\sim_B^D ww'$ we have that $\mathfrak{M}, w' \models \varphi$.

Example 4.14 ($E_B\varphi$)

E_B has been defined as macro. We show that: $\mathfrak{M}, w \models E_B\varphi$ iff for all worlds $w' \in W$ with $\sim_B^E ww'$ such that $\mathfrak{M}, w' \models \varphi$.
$\rightsquigarrow$ Blackboard

Example 4.15 (Consecutive Numbers Ctd.)

We continue Example 4.3.



$\rightsquigarrow$ Blackboard

- $\mathfrak{M}, \langle 3,2\rangle \models \neg \mathsf{b}_4 \wedge \hat{E}_{\{a,b\}}\mathsf{b}_4$
- $\mathfrak{M}, \langle 3,2\rangle \models E_{\{a,b\}}\neg \mathsf{a}_5 \wedge \neg E_{\{a,b\}}E_{\{a,b\}}\neg \mathsf{a}_5$
- $\mathfrak{M}, \langle 3,2\rangle \models E_{\{a,b\}}E_{\{a,b\}}\neg \mathsf{b}_6 \wedge \neg E_{\{a,b\}}E_{\{a,b\}}E_{\{a,b\}}\neg \mathsf{b}_6$
- $\mathfrak{M}, \langle 3,2\rangle \models \neg C_{\{a,b\}}\neg \mathsf{a}_{123} \wedge C_{\{a,b\}}\neg \mathsf{b}_{123}$

4.2 Public Announcement Language

Introduction

Example 4.16 (Public Announcement)

Consider two stockbrokers a and b sitting together. a gets a message: Company A did very well!

a says to b: **“Guess you don’t know it yet, but A is doing well.”**

a says:

- **“A is doing well”**
- **“b does not know that A is doing well”**

After the announcement b now knows that A is doing well.

- **“b does not know that A is doing well”** is now false.

- In other words: a has announced something which becomes false because of the announcement. This is called an unsuccessful update.
- Announcements refer to a specific moment in time, to a specific information state that may change because of the announcement that makes an observation about it.

p := "A is doing well"

0 := is the name of the state where p is false

1 := is the name of the state where p is true

- We assume that a only makes truthful, public announcements:

  Truthful: The formula of the announcement must be true in the actual state while the statement is made.

  Public: b can hear what a is saying, that a knows that b can hear her, etc., $\ldots$. It is common knowledge that a is making the announcement.

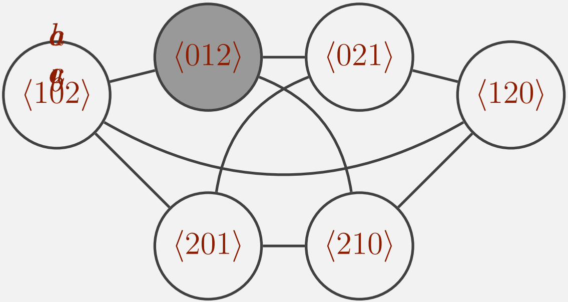
- From **truthful** and **public** it follows that states where the announcement formula is false are excluded.

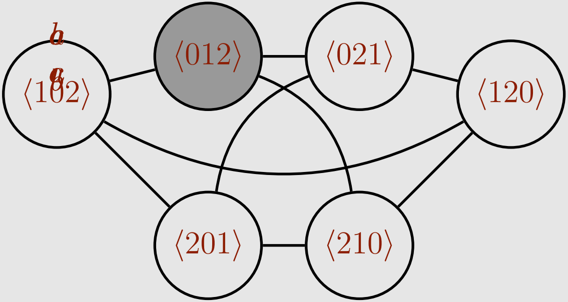
1. $\neg K_b$p
2. **Announcement**: p $\land \neg K_b$p
3. $C_{\{a,b\}}$p
4. K_bp

(1)

Example 4.17 (Three Player Card Game)

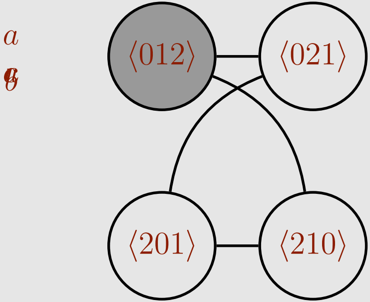
a, b and c have each drawn a card from a stack of three cards $\{0, 1, 2\}$. This is commonly known.
$V(\mathsf{a}_i) = \{\langle ijk\rangle \mid j, k \neq i, j \neq k\}$,
$V(\mathsf{b}_j) = \{\langle ijk\rangle \mid i, k \neq j, i \neq k\}$,
$V(\mathsf{c}_k) = \{\langle ijk\rangle \mid i, j \neq k, i \neq j\}$ where a_i expresses that a holds the card i etc. Assume the actual world $w = \langle 012\rangle$.



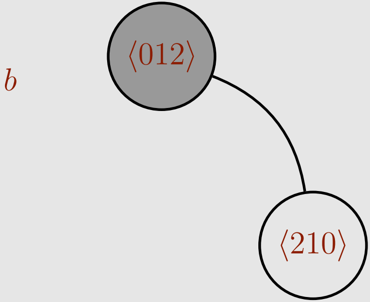


$\mathfrak{M}, \langle 012 \rangle \models K_a \neg (K_b \mathsf{a}_0 \vee K_b \mathsf{a}_1 \vee K_b \mathsf{a}_2)$ a knows that b does not know her card.

$\mathfrak{M}, \langle 012 \rangle \models \mathsf{b}_1 \wedge \hat{K}_a \mathsf{b}_2$ a considers it possible that b holds card 2 although b_1 is true.



1. Announcement of a **"I do not have card 1"**: $\neg \mathsf{a}_1$.
   - $\mathfrak{M}, \langle 012 \rangle \models K_c \mathsf{a}_0 \wedge \neg K_a K_c \mathsf{a}_0 \wedge \neg (K_b \mathsf{a}_0 \vee K_b \mathsf{a}_1 \vee K_b \mathsf{a}_2)$



2. Announcement of b **“I do not know the card of a”**:
$\neg(K_b\mathsf{a}_0 \vee K_b\mathsf{a}_1 \vee K_b\mathsf{a}_2)$
   - $\mathfrak{M}, \langle 012\rangle \models \neg(K_b\mathsf{a}_0 \vee K_b\mathsf{a}_1 \vee K_b\mathsf{a}_2)$

3. Announcement of a **“the card deal is 012”**: $a_0 \wedge b_1 \wedge c_2$
   - $\mathfrak{M} \models a_0 \wedge b_1 \wedge c_2$

Syntax

Definition 4.18 (Public Announcements $\mathcal{L}_{CPA}$)

Let $\mathcal{P}rop$ be a set of propositions and $\mathcal{A}g$ a finite set of (names for) agents. The epistemic language with public announcements and common knowledge $\mathcal{L}_{CPA}(\mathcal{P}rop, \mathcal{A}g)$ consists of all formulae defined by the following grammar:

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid C_B\varphi \mid [\varphi]\varphi$$

where $\mathsf{p} \in \mathcal{P}rop$ and $B \subseteq \mathcal{A}g$. The sublanguage without common knowledge, denoted $\mathcal{L}_{PA}(\mathcal{P}rop, \mathcal{A}g)$, is given by:

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid K_i\varphi \mid [\varphi]\varphi$$

As before, the dual operator is defined as follows:

$$\langle\psi\rangle\varphi = \neg[\psi]\neg\varphi.$$

The intuitive reading is:

$[\psi]\varphi$: After truthfully and publicly announcing ψ, φ holds.

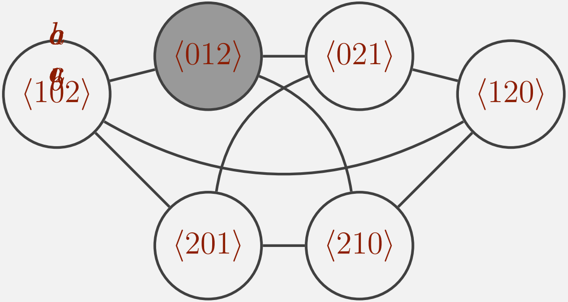
$\langle\psi\rangle\varphi$: ψ holds and after truthfully and publicly announcing ψ, φ holds.

What is the relation to $\Box$ and $\Diamond$?

If φ cannot be truthfully announced (i.e. ψ is false in the current state) formula $[\psi]\varphi$ is true $\leftrightsquigarrow$ $\Box\varphi$ is true in states with no successors.

If φ cannot be truthfully announced (i.e. ψ is false in the current state) formula $\langle\varphi\rangle\varphi$ is false $\leftrightsquigarrow$ $\Diamond\ \varphi$ is false in states with no successors.

Example 4.19 (Three Player Card Game)



1. Announcement of a **“I do not have card 1”**.
2. Announcement of b **“I do not know the card of a”**.

3a. Announcement of a **“the card deal is 012”**.

3b. Alternative announcement of a **“I now know b’s card”**:
$K_a \mathsf{b}_0 \vee K_a \mathsf{b}_1 \vee K_a \mathsf{b}_2$

b knows a’s card:

$$[\neg \mathsf{a}_1][\neg(K_b \mathsf{a}_0 \vee K_b \mathsf{a}_1 \vee K_b \mathsf{a}_2)][\mathsf{a}_0 \wedge \mathsf{b}_1 \wedge \mathsf{c}_2](K_b \mathsf{a}_0 \vee K_b \mathsf{a}_1 \vee K_b \mathsf{a}_2)$$

b does not know a’s card:

$$[\neg \mathsf{a}_1][\neg(K_a \mathsf{b}_0 \vee K_a \mathsf{b}_1 \vee K_a \mathsf{b}_2)][K_a \mathsf{b}_0 \vee K_a \mathsf{b}_1 \vee K_a \mathsf{b}_2]\neg(K_b \mathsf{a}_0 \vee K_b \mathsf{a}_1 \vee K_b \mathsf{a}_2)$$

Semantics

Definition 4.20 (Semantics $\mathfrak{M}, w \models [\psi]\varphi$)

Let $\mathfrak{M} = (W, \{\sim_i \mid i \in \mathcal{A}g\}, V)$ be a Kripke model, $w \in W_{\mathfrak{M}}$, and $\varphi \in \mathcal{L}_{CPA}$. φ is said to be **true** or **satisfied in $\mathfrak{M}$ and world w**, written as $\mathfrak{M}, w \models \varphi$, if the semantics of $\mathcal{L}_C$ is extended by the following clause:

$\mathfrak{M}, w \models [\psi]\varphi$ iff $\mathfrak{M}, w \models \psi$ **implies** $\mathfrak{M}|\psi, w \models \varphi$

where $\mathfrak{M}|\psi = (W', \{\sim'_i \mid i \in \mathcal{A}g\}, V')$:

$$\begin{aligned}
[[\psi]]_{\mathfrak{M}} &:= \{w \in W \mid \mathfrak{M}, w \models \psi\} \\
W' &:= [[\psi]]_{\mathfrak{M}} \\
\sim'_i &:= \sim_i \cap ([[\psi]]_{\mathfrak{M}} \times [[\psi]]_{\mathfrak{M}}) \\
V' &:= V \cap [[\psi]]_{\mathfrak{M}}
\end{aligned}$$

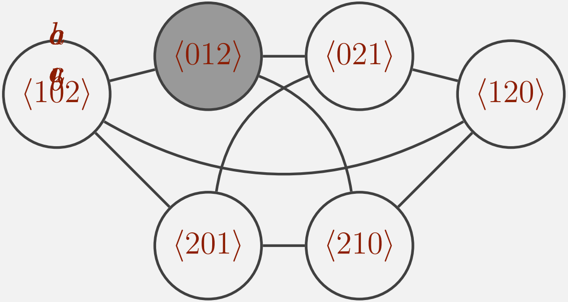
$\rightsquigarrow$ Blackboard

Remark 4.21

The semantics for the dual is defined as follows:

$\mathfrak{M}, w \models \langle\psi\rangle\varphi$ *iff* $\mathfrak{M}, w \models \psi$ **and** $\mathfrak{M}|\psi, w \models \varphi$.

Example 4.22 (Three Player Card Game)



1. Announcement of a **“I do not have card 1”**
   $\mathfrak{M}, \langle 012 \rangle \models [\neg \mathsf{a}_1] K_c \mathsf{a}_0$

We prove this property. $\mathfrak{M}, \langle 012 \rangle \models [\neg \mathsf{a}_1] K_c \mathsf{a}_0$ iff $(\mathfrak{M}, \langle 012 \rangle \models \neg \mathsf{a}_1$ implies $\mathfrak{M}|\neg \mathsf{a}_1, \langle 012 \rangle \models K_c \mathsf{a}_0)$

- Consider the former. $\mathfrak{M}, \langle 012 \rangle \models \neg \mathsf{a}_1$ iff $\mathfrak{M}, \langle 012 \rangle \not\models \mathsf{a}_1$. Due to $\langle 012 \rangle \notin V(\mathsf{a}_1) = \{\langle 102 \rangle, \langle 120 \rangle\}$ this is true.
- Consider the latter. $\mathfrak{M}|\neg \mathsf{a}_1, \langle 012 \rangle \models K_c \mathsf{a}_0$ is equivalent to $\forall w \in W_{\mathfrak{M}|\neg \mathsf{a}_1}$: $\langle 012 \rangle \sim_c w$ implies $\mathfrak{M}|\neg \mathsf{a}_1, w \models \mathsf{a}_0$. Because of the fact that only $\langle 012 \rangle$ is c-accessible from $\langle 012 \rangle$ and $\langle 012 \rangle \in V(\mathsf{a}_0) = \{\langle 012 \rangle, \langle 021 \rangle\}$ the condition is fulfilled.

4.3 Properties of Public Accouncements

Properties of Public Announcements

Most of the following properties will be used in the axiomatization of public announcement logic.

Proposition 4.23 (Negation)

$\models [\psi]\neg\varphi \leftrightarrow (\psi \rightarrow \neg[\psi]\varphi)$ *where* $\varphi, \psi \in \mathcal{L}_{CPA}$.

$[\psi]\neg\varphi$ can be true for two reasons:

1. the formula ψ cannot be announced.
2. ψ is true and after the announcement of ψ, the formula φ is false.

Proof.

$\rightsquigarrow$ Blackboard □

Proposition 4.24

The following formulae are equivalent over all ***S5****-models where* $\varphi, \psi \in \mathcal{L}_{CPA}$*:*

- $\psi \rightarrow [\psi]\varphi$
- $\psi \rightarrow \langle\psi\rangle\varphi$
- $[\psi]\varphi$

Proof.

We show: $\psi \to [\psi]\varphi$ iff $[\psi]\varphi$

$$
\begin{array}{ll}
 & \mathfrak{M}, w \models \psi \to [\psi]\varphi \\
\Leftrightarrow & \mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}, w \models [\psi]\varphi \\
\Leftrightarrow & \mathfrak{M}, w \models \psi \text{ implies } (\mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}|\psi, w \models \varphi) \\
\Leftrightarrow & (\mathfrak{M}, w \models \psi \text{ and } \mathfrak{M}, w \models \psi) \text{ implies } \mathfrak{M}|\psi, w \models \varphi \\
\Leftrightarrow & \mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}|\psi, w \models \varphi \\
\Leftrightarrow & \mathfrak{M}, w \models [\psi]\varphi
\end{array}
$$

The other points are left as exercise. □

The following properties are easy to verify:

Proposition 4.25

The following formulae are equivalent over all ***S5C****-models where* $\varphi, \psi \in \mathcal{L}_{CPA}$*:*

- $\langle\psi\rangle\varphi$
- $\psi \wedge \langle\psi\rangle\varphi$
- $\psi \wedge [\psi]\varphi$

Proof.

$\rightsquigarrow$ Exercise □

The following theorem allows to reduce the number of announcements:

Proposition 4.26 (Composition)

$[\psi \land [\psi]\varphi]\chi$ *is equivalent to* $[\psi][\varphi]\chi$ *for* $\varphi, \psi \in \mathcal{L}_{CPA}$.

Proof.

For arbitrary $\mathfrak{M}$ and w:

$$
\begin{array}{ll}
 & w \in \mathfrak{M}|(\psi \land [\psi]\varphi) \\
\Leftrightarrow & \mathfrak{M}, w \models \psi \land [\psi]\varphi \\
\Leftrightarrow & \mathfrak{M}, w \models \psi \text{ and } (\mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}|\psi, w \models \varphi) \\
\Leftrightarrow & w \in \mathfrak{M}|\psi \text{ and } \mathfrak{M}|\psi, w \models \varphi \\
\Leftrightarrow & w \in (\mathfrak{M}|\psi)|\varphi
\end{array}
$$

□

But how does the knowledge change after an announcement? Firstly, we analyse individual knowledge.

Proposition 4.27 ($[\psi]K_a\varphi \not\Leftrightarrow K_a[\psi]\varphi$)

$[\psi]K_a\varphi$ *is in general not equivalent to* $K_a[\psi]\varphi$ for $\varphi, \psi \in \mathcal{L}_{CPA}$.

Proof.

Counterexample. Consider the Three Player Card Game-Example ($\rightsquigarrow$ Blackboard):

$$\mathfrak{M}, \langle 012 \rangle \models [\mathsf{a}_1]K_c\mathsf{a}_0$$

(because the announcement cannot take place in $\langle 012 \rangle$)

but

$$\mathfrak{M}, \langle 012 \rangle \not\models K_c[\mathsf{a}_1]\mathsf{a}_0$$

□

Proposition 4.28 (Knowledge and Accouncements)

$[\psi]K_a\varphi$ is equivalent to $\psi \to K_a[\psi]\varphi$ for $\varphi, \psi \in \mathcal{L}_{CPA}$.

Proof.

For arbitrary $\mathfrak{M}$ and w:

$$\begin{aligned}
& \mathfrak{M}, w \models \psi \to K_a[\psi]\varphi \\
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}, w \models K_a[\psi]\varphi \\
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies} \\
& (\forall w' \in \mathfrak{M} : w \sim_a w' \text{ implies } \mathfrak{M}, w' \models [\psi]\varphi) \\
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies } (\forall w' \in \mathfrak{M} : w \sim_a w' \\
& \text{ implies } (\mathfrak{M}, w' \models \psi \text{ implies } \mathfrak{M}|\psi, w' \models \varphi))
\end{aligned}$$

Proof. (Cont.)

$$
\begin{aligned}
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies } (\forall w' \in \mathfrak{M} : \mathfrak{M}, w' \models \psi \text{ and } \\
& w {\sim_a} w' \text{ implies } \mathfrak{M}|\psi, w' \models \varphi) \\
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies } (\forall w' \in \mathfrak{M}|\psi, \\
& w {\sim_a} w' \text{ implies } \mathfrak{M}|\psi, w' \models \varphi) \\
\Leftrightarrow \quad & \mathfrak{M}, w \models \psi \text{ implies } \mathfrak{M}|\psi, w \models K_a\varphi) \\
\Leftrightarrow \quad & \mathfrak{M}, w \models [\psi]K_a\varphi
\end{aligned}
$$

□

Is public announcement logic without group knowledge more expressive than epistemic logic? No! We can remove all announcements! Why do we consider the logic at all?

Proposition 4.29 (From PAL to Epistemic Logic)

For $\varphi, \psi, \chi \in \mathcal{L}_{CPA}$ *we have the following equivalences:*

$$
\begin{array}{rcl}
[\psi]\mathsf{p} & \leftrightarrow & (\psi \rightarrow \mathsf{p}), \\
[\psi](\chi \vee \varphi) & \leftrightarrow & ([\psi]\chi \vee [\psi]\varphi), \\
[\psi](\chi \rightarrow \varphi) & \leftrightarrow & ([\psi]\chi \rightarrow [\psi]\varphi), \\
[\psi]\neg\varphi & \leftrightarrow & (\psi \rightarrow \neg[\psi]\varphi), \\
[\psi]K_a\varphi & \leftrightarrow & (\varphi \rightarrow K_a[\psi]\varphi), \\
[\psi][\varphi]\chi & \leftrightarrow & [\psi \wedge [\psi]\varphi]\chi.
\end{array}
$$

These validities allow us to eliminate announcements if only individual knowledge is used.

PA and Common Knowledge

In Proposition 4.28 we have shown that

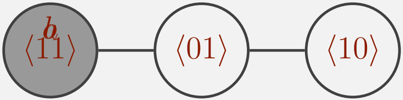
$$[\psi]K_a\varphi \leftrightarrow \psi \rightarrow K_a[\psi]\varphi$$

Is the same valid for common kowledge?

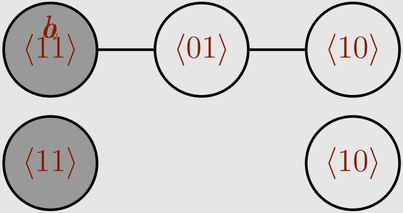
Example 4.30

Consider the following model $\mathfrak{M}$ with $V(p) = \{\langle 10\rangle, \langle 11\rangle\}$ and $V(q) = \{\langle 01\rangle, \langle 11\rangle\}$ and the instance:

$$[\mathsf{p}]C_{\{a,b\}}\mathsf{q} \leftrightarrow (\mathsf{p} \rightarrow C_{\{a,b\}}[\mathsf{p}]\mathsf{q}).$$



We show that the equivalence is false in $\langle 11\rangle$.



- $\mathfrak{M}, \langle 11 \rangle \models [\mathsf{p}] C_{\{a,b\}} \mathsf{q}$ because $\mathfrak{M}|\mathsf{p}, \langle 11 \rangle \models C_{\{a,b\}} \mathsf{q}$
- $\mathfrak{M}, \langle 11 \rangle \not\models \mathsf{p} \rightarrow C_{\{a,b\}}[\mathsf{p}]\mathsf{q}$ although $\mathfrak{M}, \langle 11 \rangle \models \mathsf{p}$ but $\mathfrak{M}, \langle 11 \rangle \not\models C_{\{a,b\}}\mathsf{q}$ because $\langle 11 \rangle \sim^{C}_{\{a,b\}} \langle 10 \rangle$ and $\mathfrak{M}, \langle 10 \rangle \not\models [\mathsf{p}]\mathsf{q}$:
  When evaluating q in $\mathfrak{M}|\mathsf{p}$ we are in its other disconnected part, where q is false: $\mathfrak{M}|\mathsf{p}, \langle 10 \rangle \not\models \mathsf{q}$

Proposition 4.31 ($[\psi]C_A\varphi \not\Leftrightarrow \psi \rightarrow C_A[\psi]\varphi$)

$[\psi]C_A\varphi$ *is in general not equivalent to* $\psi \rightarrow C_A[\psi]\varphi$ *for* $\varphi, \psi \in \mathcal{L}_{CPA}$.

The proof is given by the previous example.

We can derive common knowledge by the following rule

Proposition 4.32 (PA and CK)

Let $\varphi, \psi, \chi \in \mathcal{L}_{CPA}$. If $\chi \to [\varphi]\psi$ and $(\chi \wedge \varphi) \to E_B\chi$ are valid, then so is $\chi \to [\varphi]C_B\psi$.

Proof.

$\leadsto$ Blackboard

Let the premises $\chi \to [\varphi]\psi$ and $(\chi \wedge \varphi) \to E_B\chi$ be valid.
Suppose $\mathfrak{M}, w \models \chi$. We show that $\mathfrak{M}, w \models [\varphi]C_B\psi$.
If $\mathfrak{M}, w \not\models \varphi$ we are done. So, suppose $\mathfrak{M}, w \models \varphi$ and let $w' \in W_{\mathfrak{M}|\varphi}$ with $w \sim_B^C w'$.
By induction on the length of the path from w to w' we show $\mathfrak{M}|\varphi, w' \models \psi$.

Proof. (Cont.)

$|Path| = 0$: Then $w = w'$, and $\mathfrak{M}|\varphi, w \models \psi$ follows from $\mathfrak{M}, w \models \chi$ and $\models \chi \to [\varphi]\psi$.

$|Path| = n+1$ for $n \in \mathbb{N}$: Suppose the path has length $n+1$ where $w \sim_a w'' \sim_B^C w'$ for $a \in B$ and $w'' \in \mathfrak{M}|\varphi$. From $\mathfrak{M}, w \models \chi$, $\mathfrak{M}, w \models \varphi$, and $\models (\chi \wedge \varphi) \to E_B\chi$ it follows $\models (\chi \wedge \varphi) \to K_a\chi$ and hence $\mathfrak{M}, w'' \models \chi$ follows as $\sim_a ww''$. Because of $w'' \in \mathfrak{M}|\varphi$ we have $\mathfrak{M}, w'' \models \varphi$. We now apply the induction hypothesis and obtain $\mathfrak{M}|\varphi, w' \models \psi$.

□

Corollary 4.33

$[\varphi]\psi$ *is valid iff* $[\varphi]C_B\psi$ *is valid.*

Proof.

From right to left is obvious. From left to right follows when taking $\chi = \top$ in Proposition 4.32. □

4.4 Epistemic Logic

Axiomatization

in this section we present a sound and complete axiomatization for $\mathcal{L}_{CD}$. The part without common and distributed knowledge has already been shown to be sound and complete for **S5**. For common knowledge we proceed similarly by construction of the canonical model.

But as worlds of the model we cannot simply take the maximal consistent sets (cf. Def 3.40). Consider

$$\{E_B^n \mathsf{p} \mid n \in \mathbb{N}\} \cup \{\neg C_B \mathsf{p}\}$$

and recall the construction in Lindenbaums's theorem 3.42:

$$\Sigma_{n+1} := \begin{cases} \Sigma_n \cup \{\Phi_n\}, & \text{if this theory is consistent;} \\ \Sigma_n \cup \{\neg\Phi_n\}, & \text{else.} \end{cases}$$

and $\Sigma^+ := \bigcup_{n\geq 0} \Sigma_n$.
We notice that $\mathcal{L}_C$ is not compact!

For the common knowledge operator and the basic modal system **S5C** we have to write down all these axioms for each agent $i \in \mathcal{A}g$:

Definition 4.34 (Modal System S5C)

- K: $C_{\{i\}}(\varphi \rightarrow \psi) \rightarrow (C_{\{i\}}\varphi \rightarrow C_{\{i\}}\psi)$ (distribution of $C_{\{i\}}$ over $\rightarrow$)
- T: $C_{\{i\}}\varphi \rightarrow \varphi$ (truth)
- 4: $C_{\{i\}}\varphi \rightarrow C_{\{i\}}C_{\{i\}}\varphi$ (positive introspection)
- 5: $\neg C_{\{i\}}\varphi \rightarrow C_{\{i\}}\neg C_{\{i\}}\varphi$ (negative introspection)
- MP: $\frac{\varphi, \varphi \rightarrow \psi}{\psi}$ (modus ponens)
- NEC: $\frac{\varphi}{C_{\{i\}}\varphi}$ (necessitation of $C_{\{i\}}$)

This part is just our definition of K_i.

Axioms for everybody knowledge and common knowledge:

E: $E_B\varphi \leftrightarrow \bigwedge_{i\in B} C_{\{i\}}$ (everybody knows)

C1: $C_B(\varphi \rightarrow \psi) \rightarrow (C_B\varphi \rightarrow C_B\psi)$ (distribution of C_B over $\rightarrow$)

C2: $C_B\varphi \rightarrow (\varphi \wedge E_B C_B\varphi)$ (mix)

C3: $C_B(\varphi \rightarrow E_B\varphi) \rightarrow (\varphi \rightarrow C_B\varphi)$ (induction of common knowledge)

NECC: $\frac{\varphi}{C_B\varphi}$ (necessitation of C_B)

where $B \subseteq \mathcal{A}g$.

Note, that E_B is defined as macro!

Exercise: Show that axioms **K, D,T, 4**, and **5** for operator C_B can be derived from **S5C**.

To show completeness we proceed as follows:

1. We define the closure of a formula $\rightsquigarrow$ finite MCS's.
2. We construct canonical models wrt. the closure. These models are finite!
3. We prove Lindenbaum's theorem (over finite sets) and basic properties about the canonical model.
4. We prove the truth lemma which gives us the desired completeness result. We show:

Every S5C-consistent set is satisfiable.

Note: The sublanguage without common and distributed knowledge is shown to be complete as before. The following is only necessary due to group knowledge.

Closure

Recall the notion $Sub(\varphi)$ from Defintion 3.61.

Definition 4.35 (Closure)

We define the closure of φ, $cl(\varphi)$, as follows:

1. $\varphi \in cl(\varphi)$,
2. if $\psi \in cl(\varphi)$ then $Sub(\psi) \subseteq cl(\varphi)$,
3. if $\psi \in cl(\varphi)$ and ψ is not negated then $\neg\psi \in cl(\psi)$,
4. if $C_B\psi \in cl(\varphi)$ then $\{K_iC_B\psi \mid i \in B\} \subseteq cl(\varphi)$,

We would like to build maximal S5C-consistent sets over $cl(\varphi)$. Note that the closure is finite!

Why do we need rule 4?

Maximal Consistency

We define maximal consistency wrt. to a set of formulae.
When is a set maximal consistent?

Definition 4.36 (T-maximal S5C-consistent)

Let $T \subseteq \mathcal{L}_C$. We say that a set Γ is T-maximal S5C-consistent iff

1. $\Gamma \subseteq T$,
2. Γ is **S5C**-consistent (i.e. $\Gamma \not\vdash_{\mathbf{S5C}} \bot$),
3. Γ is a maximal set with these properties wrt. T.

The set of all T-maximal **S5C**-consistent sets is defined as $MCS_T^{\mathbf{S5C}}$.

For a finite set X of formulae we define $\underline{X}$ as $\bigwedge_{\varphi \in X} \varphi$.

Lindenbaum's theorem over finite sets is trivially true.

We define the canonical **S5C**-model completely analogously to Definition 3.43.

Definition 4.37 (Canonical Model)

The canonical model $\mathfrak{M}_{\varphi}^{\mathbf{S5C}} = (W^{\mathbf{S5C}}, \{\sim_a^{\mathbf{S5C}} \mid a \in \mathcal{A}g\}, V^{\mathbf{S5C}})$ for S5C and φ is defined as follows:

1. $W^{\mathbf{S5C}} = MCS^{\mathbf{S5C}}_{cl(\varphi)}$;
2. $\sim_a^{\mathbf{S5C}} \subseteq W^{\mathbf{S5C}} \times W^{\mathbf{S5C}}$ such that $\Gamma \sim_a^{\mathbf{S5C}} \Gamma'$ iff $\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Gamma'\}$;
3. $V^{\mathbf{S5C}}(p) := \{w \in W^{\mathbf{S5C}} \mid p \in w\}$.

$\mathfrak{F}_{\varphi}^{\mathbf{S5C}} = (W^{\mathbf{S5C}}, \{\sim_a^{\mathbf{S5C}} \mid a \in \mathcal{A}g\})$ is the canonical frame of **S5C** and φ.

What is the justification for the definition of the accessibility relations?

We need some notation to show completeness.

Definition 4.38 (Paths)

Let $B \subseteq \mathcal{A}g$, T be a set of formulae and $\Gamma \in MCS_T^{\mathsf{S5C}}$.
A (B, Γ)**-path** is given by $\Gamma_0, \ldots, \Gamma_n$ where

- each $\Gamma_i \in MCS_T^{\mathsf{S5C}}$,
- $\Gamma_0 = \Gamma$, and
- for all k with $0 \leq k < n$ there is an $a \in B$ such that $\Gamma_k \sim_a^{\mathsf{S5C}} \Gamma_{k+1}$.

A φ**-path** is given by $\Gamma_0, \ldots, \Gamma_n$ where each $\Gamma_i \in MCS_T^{\mathsf{S5C}}$ and $\varphi \in \Gamma_k$ forall $k = 0, \ldots, n$.

Such paths are closely related to the semantics of $\sim_B^C$. How?

The following is fundamental for the truth lemma.

Lemma 4.39 (Common knowledge)

Let T be a set of formulae and $\Gamma \in MCS_T^{\mathbf{S5C}}$.
If $C_B\varphi \in T$ then $C_B\varphi \in \Gamma$ iff every (B, Γ)-path is a φ-path.

Proof.

"$\Rightarrow$": We proceed by induction on the length of a path.
Suppose $C_B\varphi \in \Gamma$. We show the stronger statement every (B, Γ)-path is a φ-path and $C_B\varphi$-path.

- $n = 0$: By $\vdash C_B\varphi \rightarrow \varphi$ we have $\{\varphi, C_B\varphi\} \subseteq \Gamma$.
- Let the claim be true for paths of length n. Consider a path of length $n + 1$. We have $C_B \in \Gamma_n$ and there is an agent $a \in B$ such that $\Gamma_n \sim_a \Gamma_{n+1}$.

□

- From $\vdash C_B\varphi \rightarrow E_B C_B\varphi$ and $\vdash E_B C_B\varphi \rightarrow K_a C_B\varphi$ we have $K_a C_B\varphi \in \Gamma_n \cap \Gamma_{n+1}$ due to $\sim_a$. Again, due to the truth axiom, we have $\{\varphi, C_B\varphi\} \subseteq \Gamma_{n+1}$.

"$\Leftarrow$": Let every (B,Γ)-path be a φ-path.

- Let $G \subseteq MCS_\varphi^{\mathbf{S5C}}$ consist of all Δ such that every (B,Δ)-path is a φ-path. We define

$$\chi \equiv \bigvee_{\Delta \in G} \underline{\Delta}.$$

Suppose we have proven: (1) $\vdash \underline{\Gamma} \rightarrow \chi$, (2) $\vdash \chi \rightarrow \varphi$, and (3) $\vdash \chi \rightarrow E_B\chi$. From (3) and NECC $\vdash C_B(\chi \rightarrow E_B\chi)$. By the induction ax. $\vdash \chi \rightarrow C_B\chi$, then by (1) $\vdash \underline{\Gamma} \rightarrow C_B\chi$. Finally by (2), NECC, and the distribution ax. $\vdash \underline{\Gamma} \rightarrow C_B\varphi$ which shows that $C_B\varphi \in \Gamma$.

It remains to show the validity of (1-3). (1-2) $\rightsquigarrow$ Exercise. We show (3).

- Suppose for the sake of contradiction $\chi \wedge \neg E_B \chi$ is consistent.
- Then, there must be a disjunct $\underline{\Delta}$ of χ such that $\underline{\Delta} \wedge \neg E_B \chi$ is consistent and also an $a \in B$ such that $\underline{\Delta} \wedge \neg K_a \chi$ which is $\underline{\Delta} \wedge \hat{K}_a \neg \chi$
- From $\vdash \bigvee_{\Delta \in MCS_\varphi^{\mathsf{S5C}}} \underline{\Delta}$ (Why?) we have that $\underline{\Delta} \wedge \hat{K}_a \bigvee_{\Delta' \in MCS_\varphi^{\mathsf{S5C}} \setminus G} \underline{\Delta'}$ is consistent and by modal reasoning $\underline{\Delta} \wedge \bigvee_{\Delta' \in MCS_\varphi^{\mathsf{S5C}} \setminus G} \hat{K}_a \underline{\Delta'}$ is.
- Hence, there is $\Theta \in MCS_\varphi^{\mathsf{S5C}} \setminus G$ with $\underline{\Delta} \wedge \hat{K}_a \underline{\Theta}$ is consistent, but then $\Delta \sim_a^{\mathsf{S5C}} \Theta$ (Why?) and there is a (B, Θ)-path which is not a φ-path, but then the same holds for Δ. Contradiction to $\Delta \in G$. $\square$

Lemma 4.40 (Truth lemma)

Let $\mathfrak{M}_\varphi^{\mathsf{S5C}}$ be the canonical model for φ. Then, we have

$$\varphi \in w \quad \textbf{iff} \quad \mathfrak{M}_\varphi^{\mathsf{S5C}}, w \models \varphi.$$

Proof.

The proof is done by structural induction. We do only consider the case for $C_A\psi$.

- $C_A\psi \in w$ iff (Lemma 4.39) every (A, w)-path is a ψ path iff $\psi \in w'$ for every state w' reachable from w by $\sim_{i_1} \circ \cdots \circ \sim_{i_l}$ where $\sim_{i_j} \in \bigcup_{a \in A} \sim_a$ for $j = 1, \ldots, l$ iff $\psi \in w'$ for every state w' with $w \sim_A^C w'$ iff $\mathfrak{M}, w' \models \psi$ for every state w' with $w \sim_A^C w'$ iff $\mathfrak{M}_\varphi^{\mathsf{S5C}}, w \models C_A\psi$.

□

The following is immediate from the definition of the accessibility relation.

Proposition 4.41

The canonical model for $\varphi \in \mathcal{L}_C$ is an S5C-model.

Theorem 4.42 (Soundness and Completeness)

*System **S5C** is sound and complete for $\mathcal{L}_C$ with respect to Kripke frames $(W, \{\sim_i \mid i \in \mathcal{A}g\})$ where $\sim i$ is an equivalence relation.*

Proof.

Soundness is left as exercise. Suppose $\models \varphi$ and $\not\vdash \varphi$. Hence, $\{\neg\varphi\}$ is consistent. Then, there is an $w \in MCS^{\mathbf{S5C}}_{\neg\varphi}$ with $\neg\varphi \in w$. By the truth lemma $\mathfrak{M}^{\mathbf{S5C}}_{\neg\varphi}, w \models \neg\varphi$. Contradiction. □

Distributed Knowledge

We will present an axiomatization for distributed knowledge. The main axiom is illustrated by

$$K_i\varphi \to D_A\varphi \qquad i \in A$$

indicating that if φ is known by i then any group A containing i knows φ.

Otherwise, the operator behaves as a standard knowledge operator.

Definition 4.43 (S5CD)

The axiomatization S5CD extends **S5C** by the following axioms for distributed knowledge:

- DK: $D_B(\varphi \rightarrow \psi) \rightarrow (D_B\varphi \rightarrow D_B\psi)$ (distribution of D_B over $\rightarrow$)
- DT: $D_B\varphi \rightarrow \varphi$ (truth)
- D4: $D_B\varphi \rightarrow D_B D_B\varphi$ (positive introspection)
- D5: $\neg D_B\varphi \rightarrow D_B\neg D_B\varphi$ (negative introspection)
- NECD: $\frac{\varphi}{D_B\varphi}$ (necessitation of D_B)
- D1: $D_i\varphi \leftrightarrow K_i\varphi$
- D2: $D_B\varphi \rightarrow D_{B'}\varphi$ for $B \subseteq B'$

In order to show completeness we extend the notion of closure given in Defintion 4.35 by the following conditions:

1. $D_{\{i\}}\psi \in cl(\varphi)$ iff $K_{\{i\}}\psi \in cl(\varphi)$,
2. if $D_B\psi \in cl(\varphi)$ then $D_{B'}\psi \in cl(\varphi)$ for all $B \subseteq B' \subseteq \mathcal{A}g$.

Theorem 4.44 (Soundness and Completeness)

*System **S5CD** is sound and complete for $\mathcal{L}_{CD}$ with respect to Kripke frames $(W, \{\sim_i \mid i \in \mathcal{A}g\})$ where $\sim_i$ is an equivalence relation.*

In order to show completeness we proceed as follows:

1. We show that every S5CD-consistent set is pseudo-satisfiable.
2. For each Kripke model there is an equivalent tree-like Kripke model.
3. If φ is pseudo-satisfiable then also satisfiable.

Definition 4.45 (Pseudo-S5-model)

A pseudo-(S5-)model is an **S5**-model $(W, \{\sim_i \mid i \in \mathcal{A}g \cup \{A \mid A \subseteq \mathcal{A}g, A \neq \emptyset\}\}, V)$ with $\sim_i = \sim_{\{i\}}$.

That is, such a model simply adds "groups of agents" A. Satisfaction in pseudo models, denoted $\models^P$, is defined as before but we replace the clause for distributed knowledge by

$$\mathfrak{M}, w \models^P D_A\varphi \text{ iff } \forall w'(w \sim_A w' \text{ then } \mathfrak{M}, w' \models^P \varphi)$$

That is, we interpret D_A as "standard" knowledge operator.

We say that φ is pseudo-satisfiable if there is a pseudo-model $\mathfrak{M}$ and a state w such that $\mathfrak{M}, w \models^P \varphi$.

Lemma 4.46

Let $\varphi \in \mathcal{L}_{CD}(\mathcal{P}rop, \mathcal{A}g)$. If φ is S5CD-consistent *then it is* pseudo-satisfiable.

Proof.

We construct a canonical pseudo-model similar to the construction of the canonical $\mathfrak{M}_{\varphi}^{\mathsf{S5C}}$-model from Definition 4.37 by handling each D_A operator as a "standard" knowledge operator.

We consider S5CD-consistent sets and use the modified definition of closure. This model pseudo satisfies φ (Why?$\rightsquigarrow$ Exercise). It remains to show that $\sim_a = \sim_{\{a\}}$ to ensure that it really is a pseudo model. This is immediate due to the axiom $D_i\varphi \leftrightarrow K_i\varphi$ and due to construction (cf. Theorem 4.42). □

In the following we show that every **S5**-model is equivalent to a tree-like model. First, we need to make this notion precise.

Definition 4.47 (Path, reduced)

Let $\mathfrak{M} = (W, \{\sim_i \mid i \in \mathcal{A}g\}, V)$ be a **S5**-model and $w, w' \in W$. A (w, w')-path is given by $w_0 \sim_{i_0} \cdots \sim_{i_{k-1}} w_k$ with $w_0 = w$, $w_k = w'$ and $w_j \sim_{i_j} w_{j+1}$ for $j = 0, \ldots, k-1$.
A (w, w')-path $w_0 \sim_{i_0} \cdots \sim_{i_{k-1}} w_k$ is said to be reduced if $i_j \neq i_{j+1}$ for $j = 0, \ldots, k-1$.

Definition 4.48 (Tree like)

An **S5**-model is called tree-like if there is at most one reduced (w, w')-path between any two worlds w and w' of $\mathfrak{M}$.

Lemma 4.49

*For every **S5**-model $\mathfrak{M}$ there is a* **tree-like S5-model** *$\mathfrak{M}'$ such that for all $w \in W_{\mathfrak{M}}$ there is a state $w' \in W_{\mathfrak{M}'}$ such that for all $\varphi \in \mathcal{L}_{CD}$, $\mathfrak{M}, w \models \varphi$* **iff** *$\mathfrak{M}', w' \models \varphi$, and vice versa.*

Proof.

$\rightsquigarrow$ Exercise □

How "big" are these tree-like models? $\rightsquigarrow$ Exercise

Lemma 4.50

If φ is pseudo satisfiable *then it is also* S5-satisfiable.

proof

First, we consider the case for *one agent*, $n = 1$.

- Let $\mathfrak{M}' = (W, \sim_1, \sim_{\{1\}}, V)$ be a satisfying pseudo model and let be $\mathfrak{M} = (W, \sim_1, V)$.
- By structural induction, we show $\mathfrak{M}', w \models^P \psi$ iff $\mathfrak{M}, w \models \psi$
- All cases apart from $\psi = D_1\psi'$ are clear.
- For $\psi = D_1\psi'$ we have that $\rightsquigarrow$ Blackboard
  $\mathfrak{M}', w \models^P D_1\psi'$ iff $\mathfrak{M}, w \models K_1\psi'$ (by definition of a pseudo model and induction hypothesis) iff $\mathfrak{M}, w \models D_1\psi'$.

Now the case for $n \geq 2$. By Lemma 4.49 we can assume that the pseudo model $\mathfrak{M}'$ is **tree-like** and $\mathfrak{M}', w \models \varphi$. (The relations are denoted $\sim'$.) For each $i \in \mathcal{A}g$ we define

$$\sim_i := (\sim'_i \cup \bigcup\{\sim'_A \mid A \subseteq \mathcal{A}g, i \in A\})^+$$

Each $\sim_i$ is an equivalence relation (**Why?**). We define

$$\mathfrak{M} = (W_{\mathfrak{M}'}, \{\sim_i \mid i \in \mathcal{A}g\}, V_{\mathfrak{M}'}).$$

Now we show by structural induction on ψ that

$$\mathfrak{M}', w \models^P \psi \text{ iff } \mathfrak{M}, w \models \psi.$$

We only show the two cases $\psi = K_i\gamma$ and $\psi = D_A\gamma$.

Case: $\psi = K_i\gamma$.$\rightsquigarrow$ Blackboard

"$\Leftarrow$": Suppose $\mathfrak{M}', w \not\models^P K_i\gamma$. Hence, there is a state w' with $w \sim_i' w'$ such that $\mathfrak{M}', w' \not\models^P \gamma$. Since $\sim_i' \subseteq \sim_i$ and by induction hypothesis, $\mathfrak{M}, w \not\models K_i\gamma$.

"$\Rightarrow$": Suppose $\mathfrak{M}', w \models^P K_i\gamma$. We have to show that $\mathfrak{M}, w' \models \gamma$ for all w' with $w \sim_i w'$. If $w \sim_i w'$ then there are nodes $w = w_1, \ldots, w_k = w'$ which are interconnected by $\sim_i' \cup \bigcup_{A \subseteq \mathcal{A}g, i \in A} \sim_A'$. If $(\star)$ $\mathfrak{M}', w_j \models^P K_i\gamma$ for all $j = 1, \ldots, k$ then $\mathfrak{M}, w' \models \gamma$. Because, as $\mathfrak{M}', w' \models^P K_i\gamma$ then by axiom **T** also $\mathfrak{M}', w' \models^P \gamma$ and by induction hypothesis $\mathfrak{M}, w' \models \gamma$. It remains to show $(\star)$. Suppose the claim holds up to w_j. Then, also $\mathfrak{M}', w_j \models^P K_iK_i\gamma$ (axiom 4). If $w_j \sim_i' w_{j+1}$ then also $\mathfrak{M}', w_{j+1} \models^P K_i\gamma$. Else, if $w_j \sim_A' w_{j+1}$ we have $\mathfrak{M}', w_{j+1} \models^P K_i\gamma$ due to axioms $D_A\varphi \rightarrow D_B\varphi$, $A \subseteq B$ and $K_i \leftrightarrow D_i$.

Case: $\psi = D_A\gamma$ and $i \in A$.$\rightsquigarrow$ Blackboard
"$\Leftarrow$": Suppose $\mathfrak{M}', w \not\models^P D_A\gamma$. Then, there is a w', $w \sim'_A w'$, $\mathfrak{M}', w' \not\models^P \gamma$. By construction $\sim'_A \subseteq \sim_i$, hence $\mathfrak{M}, w' \not\models \gamma$ for each $i \in A$. This shows $\mathfrak{M}, w \not\models D_A\gamma$.
"$\Rightarrow$": Suppose $\mathfrak{M}, w \not\models D_A\gamma$.

- Then, $\exists w'(w \sim_A^D w'$ and $\mathfrak{M}, w' \not\models \gamma)$.
- Then, $\exists w' \forall a \in A(w \sim_a w'$ and $\mathfrak{M}, w' \not\models \gamma)$.
- Then, $\exists w' \forall a \in A(w \sim_a w'$ and $\mathfrak{M}', w' \not\models^P \gamma)$ by induction hypothesis. Then,
- $\exists w' \forall a \in A \exists$ reduced (w, w')-path $\lambda_a = w_0^a \sim_{i_0^a} \cdots \sim_{i_{k-1}^a} w_k^a$ with $i_j^a \in \{a\} \cup \{A \mid a \in A\}$ and $\mathfrak{M}', w' \not\models^P \gamma$,
- but then $\lambda_a = \lambda_b$ for all $a, b \in A$, and furthermore each $i_j^a = A$ (Why?).
- Hence, $k = 1$, $w \sim'_A w'$, $\mathfrak{M}', w' \not\models^P \gamma$ and thus $\mathfrak{M}', w \not\models^P D_A\gamma$.

Finally we can state the sound-and completeness proof.

of Theorem 4.44.

Let φ be an S5CD-consistent formula. Then it is pseudo satisfiable (Lemma 4.46) and S5-satisfiable (Lemma 4.50).
As before, suppose $\models \varphi$ and $\not\vdash \varphi$. Then, $\neg\varphi$ is **S5CD**-consistent and thus **S5**-satisfiable. Contradiction! □

4.5 Public Announcement Logic

Announcements without Common Knowledge

The *axiomatization for public announcement logic without common knowledge* is an extension of system **S5**.

Definition 4.51 (System PA)

System PA extends **S5** by the following axioms:

$$
\begin{array}{rl}
[\varphi]\mathsf{p} \leftrightarrow (\varphi \rightarrow \mathsf{p}) & \text{(atomic permanence)} \\
[\varphi]\neg\psi \leftrightarrow (\varphi \rightarrow \neg[\varphi]\psi) & \text{(PA and negation)} \\
[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi) & \text{(PA and conjunction)} \\
[\varphi][\psi]\chi \leftrightarrow ([\varphi \wedge [\varphi]\psi]\chi) & \text{(PA composition)} \\
[\varphi]K_i\psi \leftrightarrow (\varphi \rightarrow K_i[\varphi]\psi) & \text{(PA and knowledge)}
\end{array}
$$

Soundness is shown above.

Here, we follow a different approach to show completeness, not directly via canonical models! Thanks to Proposition 4.29 we can translate all axioms of **PA** to formulae of $\mathcal{L}_K$ (without public announcements!).

Suppose we have defined the translation $tr : \mathcal{L}_{PA} \to \mathcal{L}_K$ such that the following lemma holds.

Lemma 4.52

For all $\varphi \in \mathcal{L}_{PA}$ *it holds that* $\vdash_{PA} \varphi \leftrightarrow tr(\varphi)$.

Then, we can use the sound and complete axiomatization for **S5** to prove sound and completeness for PA.

Theorem 4.53

PA is sound and complete for $\mathcal{L}_{PA}$ over the class of epistemic Kripke models.

Proof.

Suppose $\models \varphi$, then $\models tr(\varphi)$ by Lemma 4.52. As $tr(\varphi) \in \mathcal{L}_K$ we have $\vdash_{\mathbf{S5}} tr(\varphi)$ by completeness for **S5**. Now, from **S5** $\subseteq$ **PA** we conclude $\vdash_{\mathbf{PA}} tr(\varphi)$ and by Lemma 4.52 we get $\vdash_{\mathbf{PA}} \varphi$. □

It remains to define tr and to prove Lemma 4.52.

Definition 4.54 (Translation tr)

We define $tr : \mathcal{L}_{PA} \rightarrow \mathcal{L}_K$ as follows:

$$
\begin{aligned}
tr(\mathsf{p}) &= \mathsf{p} \\
tr(\neg\varphi) &= \neg tr(\varphi) \\
tr(\varphi \wedge \psi) &= tr(\varphi) \wedge tr(\psi) \\
tr(K_a\varphi) &= K_a tr(\varphi) \\
tr([\psi]\mathsf{p}) &= tr(\psi \rightarrow \mathsf{p}) \\
tr([\varphi]\neg\psi) &= tr(\varphi \rightarrow \neg[\varphi]\psi) \\
tr([\psi](\chi \wedge \varphi)) &= tr([\psi]\chi \wedge [\psi]\varphi) \\
tr([\psi]K_a\varphi) &= tr(\psi \rightarrow K_a[\psi]\varphi) \\
tr([\psi][\varphi]\chi) &= tr([\psi \wedge [\psi]\varphi]\chi)
\end{aligned}
$$

Note that $tr(\varphi) \in \mathcal{L}_K \rightsquigarrow$ Exercise

How can we prove Lemma 4.52: $\vdash_{\mathsf{PA}} \varphi \leftrightarrow tr(\varphi)$?

Can we simply apply structural induction as usual?

To prove a property for φ one assumes that the property has been proven for all real subformulae of φ.

But now consider $tr([\varphi]\neg\psi) = tr(\varphi \rightarrow \neg[\varphi]\psi)$. The formula $[\varphi]\psi$ is not a subformula of $[\varphi]\neg\psi$.

The subformulae of a formula φ define some order on φ; e.g. ψ is “smaller” than $\neg\psi$. We define a new order on $\mathcal{L}_{PA}$ which preserves this very order.

Definition 4.55 (Complexity measure c_{PA})

The **complexity measure** $c_{PA} : \mathcal{L}_{PA} \to \mathbb{N}$ is defined as follows:

$$
\begin{aligned}
c_{PA}(\mathsf{p}) &= 1, \\
c_{PA}(\neg\varphi) &= 1 + c_{PA}(\varphi), \\
c_{PA}(\varphi \land \psi) &= 1 + \max\{c_{PA}(\varphi), c_{PA}(\psi)\}, \\
c_{PA}(K_a\varphi) &= 1 + c_{PA}(\varphi), \\
c_{PA}([\varphi]\psi) &= (4 + c_{PA}(\varphi)) \cdot c_{PA}(\psi).
\end{aligned}
$$

It is easily seen that this complexity measure preserves the order imposed by the "subformula order".

Lemma 4.56

For all $\varphi, \psi, \chi \in \mathcal{L}_{PA}$ we have the following:

1. $c_{PA}(\psi) \geq c_{PA}(\varphi)$ *if* $c_{PA}(\varphi)$ *if* $\varphi \in subform(\psi)$,
2. $c_{PA}([\varphi]\mathsf{p}) > c_{PA}(\varphi \rightarrow \mathsf{p})$,
3. $c_{PA}([\varphi]\neg\psi) > c_{PA}(\varphi \rightarrow \neg[\varphi]\psi)$,
4. $c_{PA}([\varphi](\psi \wedge \chi) > c_{PA}([\varphi]\psi \wedge [\varphi]\chi)$,
5. $c_{PA}([\varphi]K_a\psi) > c_{PA}(\varphi \rightarrow K_a[\varphi]\psi)$,
6. $c_{PA}([\varphi][\psi]\chi) > c_{PA}([\varphi \wedge [\varphi]\psi]\chi)$.

Proof.

The proof is easily done by calculating the measure. We just give a proof for the last point: $c_{PA}([\varphi \wedge [\varphi]\psi]\chi) = (5{+}\max\{\varphi, [\varphi]\psi\}){\cdot}c_{PA}(\chi) = (5{+}(4{+}c_{PA}(\varphi)){\cdot}c_{PA}(\psi)){\cdot}c_{PA}(\chi) \leq (4 + c_{PA}(\varphi))(4 + c_{PA}(\psi)) \cdot c_{PA}(\chi) = c_{PA}([\varphi][\psi]\chi)$.

□

Exercise: Show that the lemma does not hold if the 4 in Definition 4.55 were replaced by 3.

Now, we prove $\vdash_{\mathsf{PA}} \varphi \leftrightarrow tr(\varphi)$.

Proof of Lemma 4.52

The proof is done by induction on $c := c_{PA}$.

Base case. $c(\varphi) = 1$, i.e. $\models \mathsf{p} \leftrightarrow tr(\mathsf{p})$. Trivial.

Induction hypothesis. For all φ with $c(\varphi) \leq n$: $\vdash \varphi \leftrightarrow tr(\varphi)$.

Induction step. The inductive step is very easy now. We illustrate it for $[\varphi][\psi]\chi$:

- $\vdash [\varphi][\psi]\chi$ iff $\vdash [\varphi \wedge [\varphi]\psi]\chi$ by the composition axiom iff $\vdash tr([\varphi \wedge [\varphi]\psi]\chi)$ by induction hypothesis iff $\vdash tr([\varphi][\psi]\chi)$.

The other cases follow completely analogously.

Public Announcement Logic

The *axiomatization for public announcement logic with common knowledge* is a combination of system **S5C** with **PA**.

Definition 4.57 (System PAC)

System PAC combines **PA** (Definition 4.51) and **S5C** (Definition 4.34) and the following two rules:

$$\frac{\varphi}{[\psi]\varphi} \quad \text{(necessitation of } [\psi]\text{)}$$

$$\frac{\chi \rightarrow [\varphi]\psi \text{ and } \chi \wedge \varphi \rightarrow E_B \chi}{\chi \rightarrow [\varphi] C_B \psi} \quad \text{(PA and common knowledge)}$$

Soundness of the latter rule is proven in Proposition 4.32.

Finally, we state that this axiomatization is sound and complete for public announcement logic.

The proof combines ideas previously introduced. Completeness is shown similarly to **S5C**. The main observation is the following (cf. Lemma 4.39):

> $[\varphi]C_B\psi \in \Gamma$ iff every (B, Γ)-path which also is a φ-path is also a $[\varphi]\psi$-path.

Structural induction is done via an extension of the complexity measure as introduced in Definition 4.55 by adding $c_{PA}(C_B\varphi) = 1 + c_{PA}(\varphi)$.

Theorem 4.58 (Soundness and Completeness)

*System **PAC** is sound and complete for $\mathcal{L}_{CPA}$ with respect to **S5C**-models.*

5. More About Models

Content of this Chapter

In this chapter we discuss advanced topics about models. We start with basic satisfaction invariance results and introduce the fundamental concept of bisimulations. The main result about bisimulations is that bisimular models are modally equivalent. In order to show the other direction, we introduce ultrafilters and ultrafilter extensions. The latter allows to completely characterize modally equivalent models.

In the last two subsections we discuss basic concepts of first-order model theory, in particular saturated models and ultraproducts. These notions are needed to prove van Benthem's Characterization Theorem which shows that the modal language is the *bisimulation invariant fragment* of the first-order language.

5.1 Invariance Results

We have introduced what it means for a formula to be **satisfied** or **valid** in a model. We also showed how to describe **structural properties** of frames. For example, we introduced formulae which were valid on **transitive**, **serial**, or **Euclidean** frames.

In the following we consider properties which cannot be expressed within the modal language and analyse when two models are not distinguishable by modal formulae. In

the following **we mostly assume that all models are of the same similarity type**! And restrict ourself to the basic modal language if the general case is obtained in the same way.

In this section we consider invariance properties of models. That is, we would like to modify models without changing the set of satisfiable and unsatisfiable formulae.

Firstly, we need to specify a criterion to compare different models.

How can such criteria look like?

Definition 5.1 (Theory)

Let $\mathfrak{M}$ be a Kripke model and $w \in W_{\mathfrak{M}}$. The **theory of w (with respect to $\mathfrak{M}$)**, denoted by $Th_{\mathfrak{M}}(w)$ is given by all modal formulae satisfied in $\mathfrak{M}, w$, i.e.

$$Th_{\mathfrak{M}}(w) = \{\varphi \mid \mathfrak{M}, w \models \varphi\}.$$

We will also write $Th(w)$ whenever the model is clear from context.

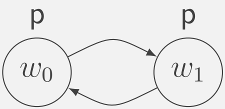
The **theory of $\mathfrak{M}$**, denoted by $Th(\mathfrak{M})$, is given by all globally valid formulae (see Definition 2.11) in $\mathfrak{M}$, i.e.

$$Th(\mathfrak{M}) = \{\varphi \mid \mathfrak{M} \models \varphi\}$$

Note: *A theory is given with respect to a set of*

Example 5.2

Consider $\mathcal{P}rop = \{p, q\}$. What is the theory of the following model?



$$\mathbf{t},\ p,\ p \wedge p, \ldots,\ p \wedge \neg q, \ldots,\ \mathbf{K},\ \mathbf{4}, \ldots,\ \neg q,\ \neg\neg\neg q, \ldots$$

Intuitively, models are considered "equal" if they have the same theory.

Definition 5.3 ((Modal) Equivalence)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two Kripke models of the same similarity type, $w_1 \in W_{\mathfrak{M}_1}$ and $w_2 \in W_{\mathfrak{M}_2}$.

We say that w_1 and w_2 are **(modally) equivalent (with respect to $\mathfrak{M}_1$ and $\mathfrak{M}_2$)**, denoted by $\mathfrak{M}_1, w_1 \leftrightsquigarrow \mathfrak{M}_2, w_2$, iff

$$Th_{\mathfrak{M}_1}(w_1) = Th_{\mathfrak{M}_2}(w_2)$$

We will also write $w_1 \leftrightsquigarrow w_2$ whenever the models are clear from context.

We say that $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are **(modally) equivalent**, denoted by $\mathfrak{M}_1 \leftrightsquigarrow \mathfrak{M}_2$ iff

$$Th(\mathfrak{M}_1) = Th(\mathfrak{M}_2)$$

In the following we describe operations on models such that the **satisfaction of formulae** is invariant under these operations. That is, if such a satisfiability preserving operation is applied to a model the resulting model should be **modally equivalent** to the original one.

We discuss the following invariance operations:

1. Disjoint unions,
2. Generated submodels,
3. Bounded morphisms,
4. Bisimulations, and
5. Ultrafilter extensions (next section).

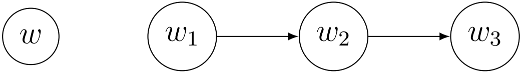
Disjoint Unions

Disjoint unions are used to construct **bigger** models from **smaller ones**.

$\mathfrak{M}_1$:

$\mathfrak{M}_2$:

$\mathfrak{M}_1 \uplus \mathfrak{M}_2$:



Definition 5.4 (Disjoint Models, - Union)

Two models $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are said to be **disjoint** iff $W_{\mathfrak{M}_1} \cap W_{\mathfrak{M}_2} = \emptyset$; i.e. if they have no worlds in common.

Let all $\mathfrak{M}_i$ be **disjoint models** for $i \in I \subseteq \mathbb{N}$. The **disjoint union of the models** $\mathfrak{M}_i$ is defined as $\biguplus_{i\in I} \mathfrak{M}_i := (W, \mathcal{R}, V)$ where

- $W := \bigcup_{i\in I} W_i$,
- $\mathcal{R} := \bigcup_{i\in I} \mathcal{R}_i$, and
- $V(\mathsf{p}) := \bigcup_{i\in I} V_i(\mathsf{p})$.

What can we say about the satisfaction in $\mathfrak{M}_i$ and $\biguplus \mathfrak{M}_i$?

Note, that **every set of models can be made disjoint** by simply renaming worlds.

Proposition 5.5 (Invariance under Disjoint U.)

Let all $\mathfrak{M}_i$ be **disjoint** *Kripke models of the same similarity type for $i \in I$. Then, for each formula φ, for each $i \in I$, and each world $w \in W_{\mathfrak{M}_i}$ it holds that*

$$\mathfrak{M}_i, w \models \varphi \quad \textit{iff} \quad \biguplus_{i \in I} \mathfrak{M}_i, w \models \varphi$$

This result is about **satisfaction**! Does it also hold for **global truth**? No! Why not?

Proof.

Structural induction on φ. $\rightsquigarrow$ Exercise □

By means of this result we can show that some operators are **not definable** in the modal language!

Example 5.6

The **global universal modality** A is defined as follows:

$$\mathfrak{M}, w \models A\phi \quad \text{iff} \quad \forall w \in W_{\mathfrak{M}} \, (\mathfrak{M}, w \models \phi)$$

Is this modality definable in the modal language?

Assume that $a(\mathsf{p})$ is a basic modal formula such that $\mathfrak{M}, w \models a(\mathsf{p})$ **iff** $\mathfrak{M}, w \models A\mathsf{p}$. That is, $a(\mathsf{p})$ **defines** A.
Let $\mathfrak{M}_1$ be a model with $V_{\mathfrak{M}_1}(\mathsf{p}) = W_{\mathfrak{M}_1}$ and $\mathfrak{M}_2$ be a model with $V_{\mathfrak{M}_2}(\mathsf{p}) = \emptyset$. Since $\mathfrak{M}_1, w \models a(\mathsf{p})$ also $\mathfrak{M}_1 \uplus \mathfrak{M}_2, w \models a(\mathsf{p})$ and thus also $\mathfrak{M}_2, w \models a(\mathsf{p})$ by Proposition 5.5.
Contradiction! So, $a(\mathsf{p})$ is **not** definable in ML.

Generated Submodels

Now, we present a way to obtain **smaller** models from **bigger** ones.

The idea is to *remove points* from a model without affecting the satisfiability of formulae.

Example 5.7

Consider the following frames: $\mathfrak{F}_1 = (\mathbb{Z}, <)$, $\mathfrak{F}_2 = (\mathbb{Z}^-, <)$ and $\mathfrak{F}_3 = (\mathbb{N}, <)$ where $\mathbb{Z}^- := \{x \in \mathbb{Z} \mid x \leq 0\}$

Is there a modal formula which can **distinguish** $\mathfrak{F}_1$ from $\mathfrak{F}_2$? Yes, consider $\Diamond \top$!

What about a formula which can **distinguish** $\mathfrak{F}_1$ from $\mathfrak{F}_3$? Why does such a formula **not** exist?

Definition 5.8 (Submodel)

Let $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ and $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ be two Kripke models. $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ is called a **submodel** of $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ iff

1. $W_{\mathfrak{M}_2} \subseteq W_{\mathfrak{M}_1}$,
2. $\mathcal{R}_{\mathfrak{M}_2} = \mathcal{R}_{\mathfrak{M}_1} \cap (W_{\mathfrak{M}_2} \times W_{\mathfrak{M}_2})$
   (i.e. $\mathcal{R}_{\mathfrak{M}_2}$ is the restriction of $\mathcal{R}_{\mathfrak{M}_1}$ to $W_{\mathfrak{M}_2}$),
3. $V_{\mathfrak{M}_2}(p) = V_{\mathfrak{M}_1}(p) \cap W_{\mathfrak{M}_2}$
   (i.e. $V_{\mathfrak{M}_2}$ is the restriction of $V_{\mathfrak{M}_1}$ to $W_{\mathfrak{M}_2}$).

That is, some points are removed and the accessibility relations and the valuation function are restricted accordingly (cf. with the modal update of public announcements).

In the following we formalize the observation made in the previous example.

Definition 5.9 (Generated Submodel)

Let $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ and $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ be two Kripke models. $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ is called a *generated submodel* of $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$, denoted by $\mathfrak{M}_2 \rightarrowtail \mathfrak{M}_1$, iff

1. $\mathfrak{M}_2$ is a **submodel** of $\mathfrak{M}_1$ and
2. if $w_2 \in W_{\mathfrak{M}_2}$ and $w_2 \mathcal{R}_{\mathfrak{M}_1} w_1$ for some $w_1 \in W_{\mathfrak{M}_1}$ then also $w_1 \in W_{\mathfrak{M}_2}$.

Let $X \subseteq W_{\mathfrak{M}_1}$. The *submodel generated by X* is defined as the **smallest** generated submodel of $\mathfrak{M}_1$ whose set of worlds **contains** X.

$\rightsquigarrow$ Exercise: Prove that the "submodel generated by X" does always exist!

If a point w is identified in some model then the submodel generated by w must contain all points reachable from w!

Example 5.10

Recall the frames from the previous example extended by some valuation: $\mathfrak{M}_1 = (\mathbb{Z}, <, V_1)$, $\mathfrak{M}_2 = (\mathbb{Z}^-, <, V_1)$ and $\mathfrak{M}_3 = (\mathbb{N}, <, V_1)$.

It is easy to see that $\mathfrak{M}_3$ is the $\{1\}$-*generated submodel of* $\mathfrak{M}_1$ but $\mathfrak{M}_2 \not\rightarrowtail \mathfrak{M}_1$.

Proposition 5.11 (Invariance under G.S.)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two models such that $\mathfrak{M}_2$ is a **generated submodel** *of $\mathfrak{M}_1$. Then, for each modal formula φ and each world $w_2 \in W_{\mathfrak{M}_2}$ it holds that*

$$\mathfrak{M}_1, w_2 \models \varphi \quad \textit{iff} \quad \mathfrak{M}_2, w_2 \models \varphi$$

Proof.

By structural induction on $\varphi \rightsquigarrow$ Exercise □

This result is quite obvious: All points reachable from w_2 are contained in both models.

How does the generated submodel look like for a model in which all points are reachable from each other?

Remark 5.12

Note that Proposition 5.5 (disjoint union) is a special case of Proposition 5.11. Why? $\rightsquigarrow$ *Exercise*

Bounded Morphisms

In the previous section we showed that models can be combined and that worlds can "carefully" be removed without affecting the satisfaction of formulae.

Now, we present a more general tool: Functions that preserve the structure of models.

Homomorphisms

We define the basic mathematical notion of a **homomorphism** for models. It is simply a mapping between states respecting relational properties and valuations.

Definition 5.13 (Homomorphism)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two models of the same similarity type. A homomorphism f from $\mathfrak{M}_1$ to $\mathfrak{M}_2$, $f : \mathfrak{M}_1 \to \mathfrak{M}_2$, is a function $f : W_{\mathfrak{M}_1} \to W_{\mathfrak{M}_2}$ with the following properties:

1. $\forall p \in \mathcal{P}rop\ \forall w \in W_{\mathfrak{M}_1}\ (w \in V_{\mathfrak{M}_1}(p) \Rightarrow \mathbf{f}(\mathbf{w}) \in \mathbf{V}_{\mathfrak{M}_2}(\mathbf{p}))$
2. $\forall w, w' \in W_{\mathfrak{M}_1}\ (\mathcal{R}_{\mathfrak{M}_1} ww' \Rightarrow \mathcal{R}_{\mathfrak{M}_2}\mathbf{f}(\mathbf{w})\mathbf{f}(\mathbf{w}'))$ (**homomorphic condition**)

Does a **homomorphism** ensure **invariance under satisfiability**?

No! The structure of $\mathfrak{M}_2$ is not reflected back to $\mathfrak{M}_1$.

$\leadsto$ Exercise: Give a counterexample.

Isomorphisms

We have to strengthen the definition of homomorphisms in order to obtain modal equivalence.

Definition 5.14 (Strong Homomorphism)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two models of the same similarity type. A **strong homomorphism** f from $\mathfrak{M}_1$ to $\mathfrak{M}_2$ is a **homomorphism** $f : \mathfrak{M}_1 \to \mathfrak{M}_2$ with the following properties:

1. $\forall p \in \mathcal{P}rop\ \forall w \in W_{\mathfrak{M}_1}\ (w \in V_{\mathfrak{M}_1}(p) \Leftrightarrow f(w) \in V_{\mathfrak{M}_2}(p))$
2. $\forall w, w' \in W_{\mathfrak{M}_1}\ (\mathcal{R}_{\mathfrak{M}_1} ww' \Leftrightarrow \mathcal{R}_{\mathfrak{M}_2} \mathbf{f}(\mathbf{w})\mathbf{f}(\mathbf{w}'))$
   (**strong homomorphic condition**)

Definition 5.15 (Isomorphism)

An **isomorphism** is a **bijective** strong homomorphism. We say that two models $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are **isomorphic**, $\mathfrak{M}_1 \backsimeq \mathfrak{M}_2$, if there is an **isomorphism** from $\mathfrak{M}_1$ to $\mathfrak{M}_2$.

Proposition 5.16 (Invariance)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two models of the same similarity type.

1. *Let w (resp. w_2) of $\mathfrak{M}_1$ (resp. $\mathfrak{M}_2$).* **If** *there is a* **surjective strong homomorphism** $f : \mathfrak{M}_1 \to \mathfrak{M}_2$ *with* $f(w_1) = w_2$ **then** $\mathfrak{M}_1, w_1 \leftrightsquigarrow \mathfrak{M}_2, w_2$.
2. *If* $\mathfrak{M}_1 \backsimeq \mathfrak{M}_2$ *then* $\mathfrak{M}_1 \leftrightsquigarrow \mathfrak{M}_2$

Proof.

The first item is shown by induction on formulae ($\rightsquigarrow$ Blackboard). The second follows from it. □

These results are not very surprising since *isomorphisms* usually describe *identical* entities (apart from renaming).

It is also not surprising that we cannot strengthen the result to an equivalence: **There are morphisms which imply modal equivalence but which are not strong!** In particular, isomorphic models have the same number of worlds.

Example 5.17

Both models are modally equivalent but not isomorphic:

$\rightsquigarrow$ Exercise: Construct a model and a morphism which is not strong but which ensures invariance under satisfiability.

The following definition refines the approach.

Bounded Morphism

Definition 5.18 (Bounded Morphism)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two models of the same similarity type. A **bounded morphism** f between $\mathfrak{M}_1$ and $\mathfrak{M}_2$, $f : \mathfrak{M}_1 \xrightarrow{b} \mathfrak{M}_2$, is a **homomorphism** $f : \mathfrak{M}_1 \rightarrow \mathfrak{M}_2$ such that the following conditions are satisfied:

1. For all $w \in W_{\mathfrak{M}_1}$ it holds that w and $f(w)$ satisfy the same propositions,
2. if $\mathcal{R}_{\mathfrak{M}_1} ww'$ then also $\mathcal{R}_{\mathfrak{M}_2} f(w) f(w')$,
3. If $\mathcal{R}_{\mathfrak{M}_2} f(w) v'$ then there is a $v \in W_{\mathfrak{M}_1}$ such that $\mathcal{R}_{\mathfrak{M}_1} wv$ and $f(v) = v'$ (**back condition**).

If the underlying homomorphism is **surjective** we say that $\mathfrak{M}_2$ is a **bounded homomorphic image** of $\mathfrak{M}_1$, $\mathfrak{M}_1 \overset{b}{\twoheadrightarrow} \mathfrak{M}_2$.

Example 5.19 (B.M. in the Basic Modal Logic)

Consider the models $\mathfrak{M}_1 = (\mathbb{N}, \mathcal{R}_1, V_1)$ and $\mathfrak{M}_2 = (\{e, o\}, \mathcal{R}_2, V_2)$ where

1. $\mathcal{R}_1 mn$ iff $n = m + 1$
2. $V_1(\mathsf{p}) = \{2n \mid n \in \mathbb{N}\}$
3. $\mathcal{R}_2 = \{(e, o), (o, e)\}$
4. $V_2(\mathsf{p}) = \{e\}$

We define $f : W_{\mathfrak{M}_1} \to W_{\mathfrak{M}_2}$ as follows:

$$f(n) = \begin{cases} e & \text{if } n \text{ is even} \\ o & \text{if } n \text{ is odd} \end{cases}$$

f is a surjective bounded morphism but **not strong!**
$\leadsto$ Exercise

Proposition 5.20 (Invariance under bm)

Let $\mathfrak{M}_1$ *and* $\mathfrak{M}_2$ *be two models of the same similarity type such that* $f : \mathfrak{M}_1 \xrightarrow{b} \mathfrak{M}_2$. *Then, for each modal formula* φ *and each* $w \in W_{\mathfrak{M}_1}$ *it holds that*

$$\mathfrak{M}_1, w \models \varphi \quad \textit{iff} \quad \mathfrak{M}_2, f(w) \models \varphi.$$

Proof.

The proof is done by induction on φ. We just consider the case where $\varphi = \Diamond \psi$.

"$\Rightarrow$": Assume $\mathfrak{M}_1, w \models \Diamond \psi$. So, there is a w' with $\mathfrak{M}_1, w' \models \psi$ and $\mathcal{R}_{\mathfrak{M}_1} ww'$. Hence, also $\mathcal{R}_{\mathfrak{M}_2} f(w) f(w')$ and by induction hypothesis $\mathfrak{M}_2, f(w') \models \psi$; hence, $\mathfrak{M}_2, f(w) \models \Diamond \psi$.

"$\Leftarrow$": Assume $\mathfrak{M}_2, f(w) \models \Diamond \psi$. So, there is a v' with $\mathfrak{M}_2, v' \models \psi$ and $\mathcal{R}_{\mathfrak{M}_2} f(w) v'$. Hence, also $\mathcal{R}_{\mathfrak{M}_1} wv$ for some v with $f(v) = v'$ and by induction hypothesis $\mathfrak{M}_1, v \models \psi$; hence, $\mathfrak{M}_1, w \models \Diamond \psi$. □

Application

Often, one is interested in models with a special structure; e.g. tree-like models. These are models which have the shape of a tree: They have a **root node** and are **acyclic**.

Proposition 5.21 (Tree Model Property)

Let $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ be a $\{w\}$-generated Kripke submodel where $w \in W_1$.
Then, there is a **tree-like** *model $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ such that $\mathfrak{M}_2 \overset{b}{\twoheadrightarrow} \mathfrak{M}_1$. This means that* any satisfiable formula is satisfiable on a tree-like model.

Proof.

W_2 consists of all **finite** sequences $(w, w_1, \ldots, w_n)$ such that

$$\mathcal{R}_1 w w_1, \mathcal{R}_1 w_1 w_2, \mathcal{R}_1 w_2 w_3, \ldots, \mathcal{R}_1 w_{n-1} w_n$$

Let $v, v' \in W_2$ we define $\mathcal{R}_2$ as follows: $\mathcal{R}_2 v v'$ iff

$v = (w, w_1, \ldots, w_n),\ v' = (w, w_1, \ldots, w_n, w_{n+1})$, and $\mathcal{R}_1 w_n w_{n+1}$

We set $(w, w_1, \ldots, w_n) \in V_2(\mathsf{p})$ iff $w_n \in V_1(\mathsf{p})$.
Finally, the function $f(w, w_1, \ldots, w_n) := w_n$ is a surjective bounded morphism from $\mathfrak{M}_2$ to $\mathfrak{M}_1$ ($\rightsquigarrow$ Why? Exercise!).

□

Bisimulations

Up to now, we introduced **special kinds** of relations between models:

- Submodels: Identity relation.
- Homomorphisms: Functions.

They have two things in common:

- Related states satisfied the same propositions, and
- transitions were preserved.

Here, we introduce a **more general relation** between models such that related models are modally equivalent.

Definition 5.22 (Bisimulation)

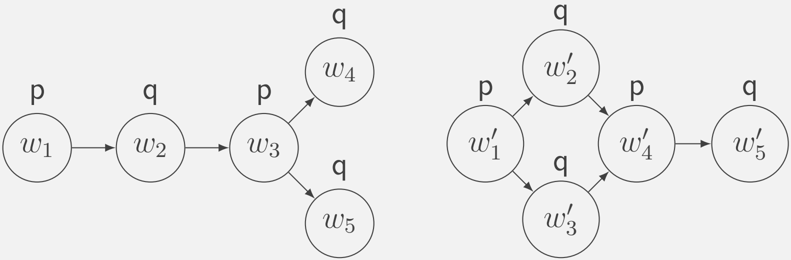
Let $\mathfrak{M}_1 = (W_1, \mathcal{R}_1, V_1)$ and $\mathfrak{M}_2 = (W_2, \mathcal{R}_2, V_2)$ be two Kripke models. A **bisimulation** between $\mathfrak{M}_1$ and $\mathfrak{M}_2$, $\mathcal{B} : \mathfrak{M}_1 \leftrightarrows \mathfrak{M}_2$, is a non-empty binary relation $\mathcal{B} \subseteq W_1 \times W_2$ such that whenever $w_1 \mathcal{B} w_2$ the following hold:

1. $V_1(w_1) = V_2(w_2)$ (**atomic harmony**)
2. if $w_1 \mathcal{R}_1 v_1$ then there is $v_2 \in W_2$ st $v_1 \mathcal{B} v_2$ and $w_2 \mathcal{R}_2 v_2$ (**force condition**)
3. if $w_2 \mathcal{R}_2 v_2$ then there is $v_1 \in W_1$ st $v_1 \mathcal{B} v_2$ and $w_1 \mathcal{R}_2 v_1$ (**back condition**)

If there is a bisimulation between two states $w_1 \in W_1$ and $w_2 \in W_2$ we write $\mathfrak{M}_1, w_1 \leftrightarrows \mathfrak{M}_2, w_2$ or just $w_1 \leftrightarrows w_2$ if the models are clear from context. If there is a bisimulation between $\mathfrak{M}_1$ and $\mathfrak{M}_2$ we write $\mathfrak{M}_1 \leftrightarrows \mathfrak{M}_2$.

Example 5.23

The following models $\mathfrak{M}$ and $\mathfrak{M}'$ are bisimular:



We define the following bisimulation between the models:
$\mathcal{B} := \{(w_1, w'_1), (w_2, w'_2), (w_2, w'_3), (w_3, w'_4), (w_4, w'_5), (w_5, w'_5)\}$.
It is easy to verify the bisimulation conditions. Note that both models are **neither a generated submodel nor a bounded homomorphic image** of the other.

How do we have to adopt the definition of bisimulation in order to fit with the general case for arbitrary similarity types?

Definition 5.24 (Bisimulation - General Case)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be τ-models. A bisimulation is defined as in Definition 5.22 but the back and force conditions are modified as follows: $w_1 \mathcal{B} w_2$ iff

- $V_1(w_1) = V_2(w_2)$ (**atomic harmony**)
- if ${\mathcal{R}_1}^m w_1 v_1 \ldots v_n$ then there are $v_1', \ldots, v_n' \in W_2$ such that $v_i \mathcal{B} v_i'$ for all $i = 1, \ldots . n$ and ${\mathcal{R}_2}^m w_2 v_1' \ldots v_n'$ (**force condition**)
- if ${\mathcal{R}_2}^m w_2 v_1' \ldots v_n'$ then there are $v_1, \ldots, v_n \in W_1$ such that $v_i \mathcal{B} v_i'$ for all $i = 1, \ldots . n$ and ${\mathcal{R}_1}^m w_1 v_1 \ldots v_n$ (**back condition**)

All constructions of the previous section are bisimulations.

Proposition 5.25

Let $\mathfrak{M}$, $\mathfrak{M}'$, and $\mathfrak{M}_i$ for $i \in I \subseteq \mathbb{N}$ be models of the same similarity type.

1. *If $\mathfrak{M} \backsimeq \mathfrak{M}'$ then $\mathfrak{M} \leftrightarrows \mathfrak{M}'$*
2. *For all $i \in I$ and $w \in W_{\mathfrak{M}_i}$ it holds that $\mathfrak{M}_i, w \leftrightarrows \biguplus_i \mathfrak{M}_i, w$*
3. *If $\mathfrak{M}' \rightarrowtail \mathfrak{M}$ then $\mathfrak{M}', w \leftrightarrows \mathfrak{M}, w$ for all $w \in W_{\mathfrak{M}'}$*
4. *If $f : \mathfrak{M} \xrightarrow{b} \mathfrak{M}'$ then $\mathfrak{M}, w \leftrightarrows \mathfrak{M}', f(w)$ for all $w \in \mathfrak{M}$.*

Proof.

1. Define $\mathcal{B}$ as $\{(w, f(w)) \mid w \in W_{\mathfrak{M}}\}$ where f is an isomorphism between $\mathfrak{M}$ and $\mathfrak{M}'$. Then, it is easy to check that $\mathcal{B}$ is a bisimulation between $\mathfrak{M}$ and $\mathfrak{M}'$.
2. Define $\mathcal{B}$ as $\{(w, w) \mid w \in W_{\mathfrak{M}_i}\}$. Then, it is easy to check that $\mathcal{B}$ is a bisimulation between $\mathfrak{M}_i$ and $\biguplus_i \mathfrak{M}_i$.
3. Define $\mathcal{B}$ as $\{(w, w) \mid w \in W_{\mathfrak{M}'}\}$. Then, it is easy to check that $\mathcal{B}$ is a bisimulation between $\mathfrak{M}'$ and $\mathfrak{M}$.
4. Define $\mathcal{B}$ as $\{(w, f(w)) \mid w \in W_{\mathfrak{M}}\}$ where f is a bounded morphism between $\mathfrak{M}$ and $\mathfrak{M}'$. Then, it is easy to check that $\mathcal{B}$ is a bisimulation between $\mathfrak{M}$ and $\mathfrak{M}'$.

□

Theorem 5.26 (Bisimulation Invariance T.)

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be models of the same similarity type. Then, for every $(w, w') \in W_{\mathfrak{M}} \times W_{\mathfrak{M}'}$ it holds that

$$w \leftrightarrows w' \textbf{ implies } \textit{that } w \leftrightsquigarrow w'.$$

The proof is done by structural induction of formulae.
What about the reverse? It does not hold in general!

Proof.

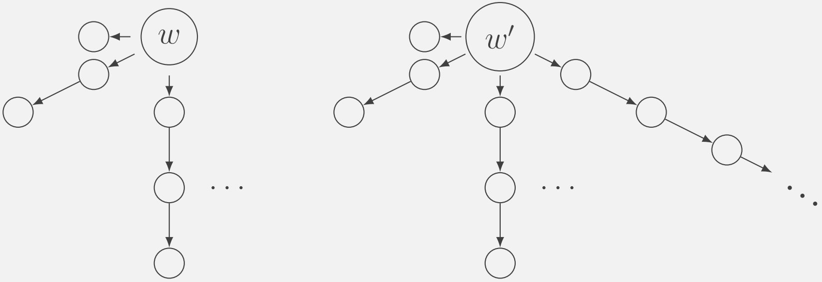
If φ is a proposition it is obvious, as well as for boolean connectives. Now let $\varphi = \Diamond\,\psi$.

Let $w \leftrightarrows w'$. Since $\mathfrak{M}, w \models \Diamond\,\psi$ we have that $\exists v(Rwv$ and $\mathfrak{M}, v \models \psi)$. Then, there is an v' $(Rw'v'$ and $v \leftrightarrows v')$ by the bisimulation force property. By induction hypothesis we get $\mathfrak{M}', v' \models \psi$. So, we have that $\mathfrak{M}', w' \models \Diamond\,\psi$.

The reverse direction is done analogously by using the back condition. □

Example 5.27

Consider the following two models:



Left model $\mathfrak{M}$: All **branches of finite length**. Right model $\mathfrak{M}'$: an additional infinite path (p is always true). Both models are **modally equivalent** in w and w'. But: there is no bisimulation!

We prove that a bisimulation cannot exist. Suppose the contrary.

Let v_1 be the state of the infinite path with $R_2 w' v_1$, and w_1 a point in $\mathfrak{M}$ bisimular with v_1. Let the path corresponding to w_1 have length $n+1$, say $ww_1 \dots w_n$. Since both models are bisimular w_n has to be bisimular with the $(n+1)$th state v_n on the infinite path as well. However, now v_n has a successor but w_n does not. Contradiction!

Definition 5.28 (Image-Finite Model)

Let $\mathfrak{M} = (W, \mathcal{R}_1, \ldots, \mathcal{R}_k, V)$ be a Kripke model of modal similarity type τ. The model $\mathfrak{M}$ is called image-finite iff for each relation $\mathcal{R}_m$ in the model the set

$$\{(x_1, \ldots, x_n) \mid \mathcal{R}_i x x_1 \ldots x_n\}$$

is **finite** where $\tau(m) = n - 1$ (i.e. $\mathcal{R}_m$ is a $n+1$-ary relation).

This definition says that the **branching factor** of each relation must be **finite**.

Theorem 5.29 (Hennessy-Milner Theorem)

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be image-finite models. Then, for every $(w, w') \in W_{\mathfrak{M}} \times W_{\mathfrak{M}'}$ it holds that

$$w \leftrightarrows w' \textbf{ if, and only if, } w \leftrightsquigarrow w'.$$

Proof.

"$\Rightarrow$": Theorem 5.26
"$\Leftarrow$": We show that $w \leftrightsquigarrow w'$ itself is a bisimulation.
Therefore, we have to verify the three properties. □

1. Clearly, w and w' satisfy the same proposition letters.
2. Forth condition: Assume Rwv and that there is **no** v' with $R'w'v'$ and $v \leftrightsquigarrow v'$. Let $S' := \{u' \mid R'w'u'\}$. The set it non-empty! Why? Image-finiteness implies that $S' = \{w'_1, \ldots w'_n\}$ is finite! By assumption there is a formula ϕ_i for each $w'_i \in S'$ such that $\mathfrak{M}, v \models \phi_i$ and $\mathfrak{M}', w'_i \not\models \phi_i$. Hence

$$\mathfrak{M}, w \models \Diamond(\phi_1 \wedge \cdots \wedge \phi_n) \quad \text{and} \quad \mathfrak{M}', w' \not\models \Diamond(\phi_1 \wedge \cdots \wedge \phi_n)$$

which contradicts $w \leftrightsquigarrow w'$.
3. Back condition: Analogously to the previous point.

□

5.2 Ultrafilter Extensions

The Picture so Far

Bisimulations are an important tool to characterize modally equivalent models. However, Example 5.27 and the Hennessy-Milner Theorem 5.29 show that bisimulations do **not** completely characterize modally equivalent models.

In this section, we introduce the final step: Ultrafilter extensions. The main result of this section is the following:

> *Two models are modally equivalent **if, and only if,** there is a bisimulation somewhere else!*

Here, "somewhere else" means in the ultrafilter extension!

$$\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w' \quad \text{iff} \quad \mathfrak{ue}(\mathfrak{M}), \pi_w \leftrightarrows \mathfrak{ue}(\mathfrak{M}'), \pi_{w'}.$$

Modal Saturation

We would like to generalize the **Hennessy-Milner** Theorem: **Image-finite** models form just a **special** class of models.

Definition 5.30 (Hennessy-Milner Property)

Let $\mathcal{M}$ be a class of τ-models. We say that $\mathfrak{M}$ has the Hennessy-Milner property iff for all $\mathfrak{M}, \mathfrak{M}' \in \mathcal{M}$, $w \in W_{\mathfrak{M}}$, $w' \in W_{\mathfrak{M}'}$ it holds that

$$w \leftrightsquigarrow w' \text{ implies } \mathfrak{M}, w \leftrightarrows \mathfrak{M}', w'.$$

This is the converse direction of Theorem 5.26.

Example 5.31

The class of image-finite models has the Hennessy-Milner property.

Definition 5.32 ((Finitely) Satisfiable)

Let $\mathfrak{M}$ be a model, $X \subseteq W_{\mathfrak{M}}$ and Σ be a set of formulae.
We say that Σ is satisfiable in X iff there is an $w \in X$ such that $\mathfrak{M}, w \models \Sigma$.
Σ is said to be finitely satisfiable in X iff every **finite** subset of Σ is satisfiable in X.

Example 5.33

Assume $\Sigma = \{\varphi_1, \varphi_2, \dots\}$ is an infinite set of formulae and that $S := \{v_1, v_2, \dots\}$ is the set of successors of a state w.
Moreover, assume that v_i satisfies exactly the formula $\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_i$ for all i. Then, Σ is **finitely satisfiable** in S but **not satisfiable**.

We mention a very important result of first-order logic and define an analogon for the modal language.

Remark 5.34 (Compactness in FOL)

Let Σ be a set of first-order formulae (or propositional formulae), possibly infinite. The compactness theorem *says that if every* **finite** *subset Σ' of Σ has a model (i.e. there is a model $\mathfrak{M}$ such that $\mathfrak{M} \models \Sigma'$) then there also is a model for Σ itself.*

In the following we present the "modal counterpart" of the compactness theorem.

The following definition generalizes the image-finite model property.

Definition 5.35 (Modal Saturation)

Let $\mathfrak{M}$ be a model. $\mathfrak{M}$ is said to be modally saturated if the following condition is satisfied for every state $w \in W_{\mathfrak{M}}$ and every set Σ of formulae:

If Σ is **finitely satisfiable** in $\{w' \in W_{\mathfrak{M}} \mid \mathcal{R}ww'\}$ then Σ itself is **satisfiable** in this set.

Proposition 5.36 (Modally Saturated Models)

The class of modally saturated models has the Hennessy-Milner property.

Proof.

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be two modally-saturated models. We show that $\leftrightsquigarrow$ is a bisimulation. We only show the **forth condition**.

Let $w, v \in W$, $w' \in W'$, Rwv, and $w \leftrightsquigarrow w'$. Let Σ be the set of formulae true at v. For every finite $\Sigma' \subseteq \Sigma$ holds: $\mathfrak{M}, v \models \bigwedge \Sigma'$, hence also $\mathfrak{M}, w \models \Diamond \bigwedge \Sigma'$. From $w \leftrightsquigarrow w'$ it follows that $\mathfrak{M}', w' \models \Diamond \bigwedge \Sigma'$, hence $\mathfrak{M}', v' \models \bigwedge \Sigma'$ for some $v' \in W'$ with $R'w'v'$. So, Σ is finitely satisfiable in $\{v' \in W' \mid R'w'v'\}$ and hence, also satisfiable in some $v' \in W'$ with $R'w'v'$. This shows that $v \leftrightsquigarrow v'$. $\square$

Ultrafilter Extensions

We have just seen that that **modally saturated** models have a very nice property. Now, we show how to **construct** such models.

An ultrafilter can be seen as a **completion** of a model: States are added to the model until it is modally saturated.

Filters and ultrafilters are just sets of sets which are **closed under** various operations.

Before we consider ultrafilter extensions we need to introduce some basic things about **filters** and **ultrafilters**.

Zorn's Lemma

For the **ultrafilter theorem** we need **Zorn's Lemma**.

Let X be a set of sets. Then, $\subseteq$ defines a **partial order** on X. X is said to be a **chain** iff for all $A, B \in X$ it holds that $A \subseteq B$ or $B \subseteq A$.

An element A of X is said to be **maximal** iff for all elements $B \in X$ with $A \subseteq B$ it follows that $A = B$.

Lemma 5.37 (Zorn's Lemma)

Let $X \neq \emptyset$ be a set of sets which is **closed under unions of non-empty chains** *over X.*
Then, X has a **maximal element.**

Filters and Ultrafilters

Definition 5.38 ((Proper) Filter)

Let $W \neq \emptyset$. A **filter F over W** is a set $F \subseteq \mathcal{P}(W)$ with the following properties:

1. $W \in F$,
2. $X, Y \in F$ implies $X \cap Y \in F$ (**closed under intersection**),
3. $X \in F$ and $X \subseteq Z \subseteq W$ implies $Z \in F$ (**upward closed**).

A filter F is called **proper** if $F \neq \mathcal{P}(W)$.

It is easy to verify that $\mathcal{P}(W)$ is a filter but no proper one.

Note that a filter is **proper** iff it does **not** contain $\emptyset$!

Definition 5.39 (Ultrafilter)

Let $W \neq \emptyset$. An **ultrafilter U over W** is a **proper filter U over W** such that the following condition holds:

$$\forall X \in \mathcal{P}(W) : X \in U \quad \textbf{iff} \quad W \backslash X \notin U$$

We use $Uf(W)$ to denote the set of ultrafilters over W.

Note the relation to a **complete theory**: Either a formula or its negation holds in it.

Example 5.40

Let $W = \{1, 2\}$. Then, $\{\{1\}, W\}$ and $\{\{2\}, W\}$ are ultrafilter over W.

Definition 5.41 (Generated Filter)

Let $W \neq \emptyset$ and $E \subseteq \mathcal{P}(W)$. The filter F generated by E is defined as follows:

$$F := \bigcap\{G \mid E \subseteq G \text{ and } G \text{ is a filter over } W\}$$

That is, F is the **smallest filter** containing E.

$\rightsquigarrow$ Exercise: Prove that F is a filter.

Definition 5.42 (Prinicpal Ultrafilter)

Let $W \neq \emptyset$ and $w \in W$. The principal ultrafilter π_w generated by w over W is the **filter generated by** $\{w\}$.

$\rightsquigarrow$ Exercise: Prove that π_w is an ultrafilter and that $\pi_w = \{X \subseteq W \mid w \in X\}$.

Definition 5.43 (Finite Intersection Property)

Let $E \subseteq \mathcal{P}(W)$. We say that E has the finite intersection property iff the **intersection** of any two elements of E is **non-empty**.

Proposition 5.44

Let F be the filter generated by $E \subseteq \mathcal{P}(W)$. F is a proper filter over W **iff** *E has the finite intersection property.*

Proof.

$\rightsquigarrow$ Blackboard □

Proof.

Firstly, we show that the set $D = \{X \subseteq \mathcal{P}(W) \mid X = W \text{ or } \exists Y_1, \ldots, Y_n \in E\ (Y_1 \cap \cdots \cap Y_n \subseteq X)\}$ is equal to F.

"$F \subseteq D$": Let $W, X, X' \in F$ and $Y_i, Y_i' \in E$ such that

$$Y_1 \cap \cdots \cap Y_n \subseteq X \text{ and } Y_1' \cap \cdots \cap Y_m' \subseteq X'$$

- $X \subseteq Z \subseteq W \Rightarrow Y_1 \cap \cdots \cap Y_n \subseteq Z \Rightarrow Z \in D$
- $Y_1 \cap \cdots \cap Y_n \cap Y_1' \cap \cdots \cap Y_m' \subseteq X \cap X' \Rightarrow X \cap X' \in D$

$\Rightarrow D$ is a filter over W. Since $E \subseteq D \Rightarrow F \subseteq D$.

"$D \subseteq F$": Let $G \subseteq \mathcal{P}(W)$ be **any** filter with $E \subseteq G$, then $W \in G$. For $Y_1, \ldots, Y_n \in E \Rightarrow Y_1 \cap \cdots \cap Y_n \in G$. So, $\forall X \in \mathcal{P}(W)\ (Y_1 \cap \cdots \cap Y_n \subseteq X \Rightarrow X \in G)$. Hence, $D \subseteq G$ and therewith also $D \subseteq F$ as F is a filter and contains E as well.

So, we have shown that F is equal to

$$\{X \subseteq \mathcal{P}(W) \mid X = W \text{ or } \exists Y_1, \ldots, Y_n \in E\ (Y_1 \cap \cdots \cap Y_n \subseteq X)\}$$

Now, it is easy to prove the proposition:

"$\Rightarrow$": Assume there are $X, X' \in E$ such that $X \cap X' = \emptyset$. Since $E \subseteq F$ also $X, X' \in F$ and hence $\emptyset = X \cap X' \in F$ and $F = \mathcal{P}(W)$.

"$\Leftarrow$": Let E has the finite intersection property. Suppose $\emptyset \in F$. Hence, there are $Y_1, \ldots, Y_n \in E$ with $Y_1 \cap \cdots \cap Y_n \subseteq \emptyset$. But this contradicts the finite intersection property of E.

Proposition 5.45

If U is a maximal proper filter *over W, then U is an* ultrafilter.

Note: The reverse direction does also hold.

Proof.

Let U be a maximal proper filter. Then not both $X \in U$ and $W \backslash X \in U$ (otherwise $\emptyset \in F$, Contradiction).

We prove: If $W \backslash X \notin U$ then $X \in U$.

Assume $W \backslash X \notin U$. Let $E := U \cup \{X\}$ and F be the filter generated by E. We show that E has thefinite intersection property. Let $Y_1, \ldots, Y_n \in E$ and $Z = \bigcap Y_i$.
Since U is closed under finite intersections it holds either

1. $Z \in U$, where $Z \neq \emptyset$ since U is proper, or
2. $Z = X \cap Y$ for $Y \in U$ where $Z \neq \emptyset$. Assume the contrary, then $Y \subseteq W \backslash X$, hence also $W \backslash X \in U$. Contradiction.

Thus, $Z \neq \emptyset$ and, by Proposition 5.44, we have $\emptyset \notin F$ (F proper). Since $U \subseteq F$ we have $U = F$, $E \subseteq U$ and $X \in U$.

□

Theorem 5.46 (Ultrafilter Theorem)

If $E \subseteq \mathcal{P}(W)$ *has the* **finite intersection property** *then* E *can be* **extended to an ultrafilter** U *over* W*; i.e.* $E \subseteq U$.

Proof.

- Let $\mathfrak{P}$ be the class of all **proper** filters over W that contain E. By Proposition 5.44 the filter generated by E is proper and thus $\mathfrak{P}$ is non-empty.
- Then, by Zorn's Lemma, $\mathfrak{P}$ **has a maximal element** U as it is closed under unions of non-empty chains: Let C be a non-empty chain of proper filters, then $\bigcup C$ is a **proper** filter. ($\rightsquigarrow$ Exercise)
- Then, $E \subseteq U$ and U is a maximal proper filter of W. For, assume M is is a proper filter in W with $U \subseteq M$, so $E \subseteq M$, hence $M \in \mathfrak{P}$ and $U = M$.
- Proposition 5.45: U is ultrafilter over W.

□

Logical Interpretation of Filters

A **proposition** can be considered as a **set of states**; namely, the states in which the proposition is true:

$$V(p) = W_p$$

Then, a filter is a (propositional) **theory** (a set of propositions) which is closed under logical operations:

$W_p \cap W_q$: Represents the conjunction. All states in which p and q are true.

$W_p \in F$ and $W_p \subseteq W_q \subseteq W$ implies $W_q \in F$: entailment "$p \rightarrow q$.

What is a **proper filter**? A **consistent** theory!

What is an **ultrafilter**? A **complete** theory!

Ultrafilter Extensions

Definition 5.47 (Pre-Image Operator)

Let $\mathcal{R}$ be a $(n+1)$-ary relation on W. Let $X_1, \ldots, X_n \in \mathcal{P}(W)$. We define the following pre-image operator:

$$pre_R(X_1, \ldots, X_n) := \{x \in W \mid \exists (x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n \text{ such that } Rxx_1 \ldots x_n\}$$

What is the interpretation of the operator for binary relations?

$$pre_R(X) = \{x \in W \mid \exists y \in X \text{ such that } Rxy\}$$

Definition 5.48 (Dual of the Pre-Image)

We define the dual of the pre-image operator for any set $X_i \subseteq W$ and $(n+1)$-ary relation R as follows:

$$pre_R^d(X_1, \ldots, X_n) := W \backslash pre(W \backslash X_1, \ldots, W \backslash X_n)$$

Note, that the **dual of the pre-image** $pre_R^d(X_1, \ldots, X_n)$ is equivalent to the following set

$$\{x \in W \mid \forall x_1, \ldots, x_n \in W \, (Rxx_1 \ldots x_n \rightarrow \exists i \, (1 \leq i \leq n \wedge x_i \in X_i))\}$$

And for the simple case:

$$pre_R^d(X) = \{x \in W \mid \forall y \in W (Rxy \rightarrow y \in X)\}$$

There is **no** *x-related element which is* **not** *in X.*

Definition 5.49 (Ultrafilter Extension)

Let $\mathfrak{F} = (W, \mathcal{R})$ be a frame. The ultrafilter extension of $\mathfrak{F}$, $\mathfrak{ue}(\mathfrak{F})$, is defined by $(Uf(W), \mathcal{R}^{\mathsf{ue}})$ where $\mathcal{R}^{\mathsf{ue}}uu'$ for $u, u' \in Uf(W)$ if the following condition holds:

If $X' \in u'$ then $pre_R(X') \in u$.

The ultrafilter extension of a model $\mathfrak{M} = (\mathfrak{F}, V)$ is defined as the model $\mathfrak{ue}(\mathfrak{M}) = (\mathfrak{ue}(\mathfrak{F}), V^{\mathsf{ue}})$ where $V^{\mathsf{ue}}(p_i) = \{u \in Uf(W) \mid V(p_i) \in u\}$.

Ultrafilter Extension: Interpretation

As mentioned above, an ultrafilter can be considered as **complete theory**, we also call it a state of affairs.

Now, not every state of affairs needs to be realized in a frame: There might **not be a single state** in a frame that satisfies **exactly** all the proposition described by the state of affairs (i.e. a state of the model is not a member of all elements of the ultrafilter).

Which states of affairs **are** realized in a frame? A **principal ultrafilter** π_w is realized (exactly) at world w of the frame!

Thus, the ultrafilter extension adds to a frame **all** states of affairs; thus, every proposition of the model is realized at some point of the ultrafilter extension.

Finally, two states are related ($\mathcal{R}^{\text{ue}}uu'$) whenever u "sees" u'.

Proposition 5.50 (Relation to Prinicpal Uf)

Let $\mathfrak{F}$ be a Kripke frame and $\mathfrak{ue}(\mathfrak{F})$ its ultrafilter extension. Then, we have for all $v, w \in W_{\mathfrak{F}}$ that

$$\mathcal{R}vw \quad \textit{iff} \quad \mathcal{R}^{ue}\pi_v\pi_w.$$

So, $\mathfrak{F}$ is isomorphic to a submodel of $\mathfrak{ue}(\mathfrak{F})$. *In general, this submodel is not a generated submodel.*

Proof.

$\rightsquigarrow$ Exercise □

Example 5.51

Consider the following frame $\mathfrak{F} = (\mathbb{N}_0, <)$.
There are two kinds of ultrafilters ($\rightsquigarrow$ Exercise):

1. **Principal** ones: They form an **isomorphic copy** of $\mathfrak{F}$ inside $\mathfrak{ue}(\mathfrak{F})$; and
2. **Non-principal** ones: They contain all **co-finite sets** and **only infinite sets**.

How are the non-principal ultrafilters related to the others?
Let $u \in Uf(\mathbb{N})$ and u' be non-principal. Then, for any $X \in u'$ we have

$$pre(X) = \{x \in \mathbb{N}_0 \mid \exists i \in X \ (x < i)\} = \mathbb{N}_{\not{0}}$$

because X is an infinite subset of $\mathbb{N}$.
Now, for any $u \in Uf(\mathbb{N}_0)$ we have that $\mathbb{N}_0 \in u$ and thus it holds that $\mathcal{R}^{\text{ue}}uu'$.

Lemma 5.52 ($\mathcal{R}^{ue}$ and pre^d)

Let $\mathfrak{M} = (W, \mathcal{R}, V)$ be given, and $u, v \in Uf(W)$. Then, we have that

$$\mathcal{R}^{ue}uv \quad \textit{iff} \quad \{Y \mid pre^d_R(Y) \in u\} \subseteq v$$

and $\{Y \mid pre^d_R(Y) \in u\}$ is **closed under finite intersection.**

Proof.

$$\mathcal{R}^{\text{ue}}uv \text{ iff } \forall X \in v\ (pre(X) \in u) \text{ iff } X \in v\ (pre^d(W \backslash X) \notin u)$$
$$\text{iff } \forall X (W \backslash X \notin v \rightarrow pre^d(W \backslash X) \notin u)$$
$$\text{iff } \forall X (pre^d(W \backslash X) \in u \rightarrow W \backslash X \in v)$$
$$\text{iff } \{Y \mid pre^d(Y) \in u\} \subseteq v$$

Now, we show that the set is closed under intersection. Assume $X, Y \in \{Y \mid pre^d(Y) \in u\}$. So, we have that $pre^d(X) \in u$ and $pre^d(Y) \in u$. Ultrafilter are closed under intersection, thus $pre^d(X) \cap pre^d(Y) \in u$. Now the set $pre^d(X) \cap pre^d(Y)$ is equivalent to

$$\{x \mid \forall y\ (Rxy \rightarrow y \in X \cap Y)\} = pre^d(X \cap Y)$$

□

Given a model $\mathfrak{M}$ and a formula φ, we write $V(\varphi)$, the denotation of φ in $\mathfrak{M}$, for the set of worlds which satisfy φ, i.e. $V(\varphi) = \{w \mid \mathfrak{M}, w \models \varphi\}$.

Proposition 5.53 (Uf Equivalence Theorem)

Let $\mathfrak{M}$ be a model, φ a formula, and $u \in Uf(W)$ an ultrafilter over W. Then it holds that

$$V(\varphi) \in u \quad \textit{iff} \quad \mathfrak{ue}(\mathfrak{M}), u \models \varphi$$

So, we have that for every state $w \in W$ it holds that $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{ue}(\mathfrak{M}), \pi_w$.

This theorem is very important as it relates **truth** in the **ultrafiler extension** with truth in the **original model** and vice versa.

Proof.

Firstly, we prove the second part: $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{ue}(\mathfrak{M}), \pi_w$.
Assume that the first part holds. Then

$$\begin{aligned}
&\mathfrak{M}, w \models \varphi \\
&\text{iff } w \in V(\varphi) \\
&\text{iff } V(\varphi) \in \pi_w \text{ (property of the principal uf)} \\
&\text{iff } \mathfrak{ue}(\mathfrak{M}), \pi_w \models \varphi \text{ (first result)}
\end{aligned}$$

Now we come to the more difficult part. We proceed by induction on φ. The cases for $\varphi = \mathrm{p}$, $\varphi = \neg\psi$ or $\varphi = \psi \wedge \psi'$ are as usual! $\rightsquigarrow$ Exercise. □

We now consider the case $\varphi = \Diamond\,\psi$.

"$\Leftarrow$":

$\mathfrak{ue}(\mathfrak{M}), u \models \Diamond\,\phi$

iff $\exists u' \in Uf(W)(\mathcal{R}^{\mathsf{ue}}uu' \wedge \mathfrak{ue}(\mathfrak{M}), u' \models \phi)$

iff $\exists u' \in Uf(W)(\mathcal{R}^{\mathsf{ue}}uu' \wedge V(\phi) \in u')$ (by induction)

iff $\exists u' \in Uf(W)(\forall X \in u'(pre(X) \in u) \wedge V(\phi) \in u')$

$\Rightarrow \exists u' \in Uf(W)(pre(V(\phi)) \in u \wedge V(\phi) \in u')$

$\Rightarrow V(\Diamond\,\phi)) \in u$ since $pre(V(\phi)) = V(\Diamond\,\phi)$

"⇒": Let $V(\Diamond\varphi) \in u$. We show $\mathfrak{ue}(\mathfrak{M}), u \models \Diamond\varphi$.
The latter is the case iff $\exists u'(\mathcal{R}^{\mathsf{ue}}uu' \wedge V(\varphi) \in u')$ and by Lemma 5.52 iff

$$\exists u'(u'_0 := \{Y \mid pre^d_R(Y) \in u\} \subseteq u' \wedge V(\varphi) \in u').$$

In the following we construct u'. We would like to apply the ultrafilter theorem.

- Show $\forall Y \in u'_0\,(Y \cap V(\varphi) \neq \emptyset)$.
  For $Y \in u'_0$ we have $pre^d(Y) \in u$ and $pre^d(Y) \cap V(\Diamond\varphi) \in u$. There must also be some element x in the intersection. $x \in V(\Diamond\varphi)$ implies that there is an y with $\mathcal{R}xy$ and $y \in V(\varphi)$. $x \in pre^d(Y)$ implies that $y \in Y$ and hence $y \in Y \cap V(\varphi)$.

- u_0' closed under intersection and $Y \cap V(\varphi) \neq \emptyset$ implies that $u_0' \cup \{V(\varphi)\}$ has the finite intersection property (cf. Def. 5.43). Hence, by the Ultrafilter Theorem 5.46 there is an ultrafilter u' with $u_0' \cup \{V(\varphi)\} \subseteq u'$.

  So we have that $\mathcal{R}^{\text{ue}}uu'$ and $V(\varphi) \in u'$ which concludes the proof.

Example 5.54 (Undefinability in $\mathcal{L}_{\mathsf{ML}}$)

We can use ultrafilter extensions to compare the **expressive power** of modal languages. Is the following operator ref definable in the modal language?

$$\mathfrak{M}, w \models ref \text{ iff } \exists w' \in W \; (Rw'w')$$

That is, this operator is true in a model which has at least one reflexive point.

Consider the frame $\mathfrak{F} = (\mathbb{N}, <)$ and its ultrafilter extension from Example 5.51.

Now, if such an operator would be definable we would **not** have $0 \leftrightsquigarrow \pi_0$ because $\mathfrak{F}$ has no loops but its ultrafilter extensions does have (infinitely many) loops. That **contradicts** the previous proposition!

Finally, we have the theorem which says that the **ultrafilter extensions** is modally saturated and therewith does have the **Hennessy-Milner property**.

Proposition 5.55 (Uf are Modally Saturated)

The ultrafilter extension $\mathfrak{ue}(\mathfrak{M})$ *of any model* $\mathfrak{M}$ *is* modally saturated.

Proof.

Let $\mathfrak{M} = (W, \mathcal{R}, V)$ be a model, $u \in Uf(W)$, and Σ a set of formulae that is finitely satisfiable in $\{v \mid \mathcal{R}^{\mathsf{ue}}uv\}$. We have to show that there is an u' in this set such that $\mathfrak{ue}(\mathfrak{M}), u' \models \Sigma$. We defined $\Delta := \{V(\varphi) \mid \varphi \in \mathsf{con}(\Sigma)\} \cup \{Y \mid pre^d(Y) \in u\}$ where **con**(Σ) is the set of **finite conjunctions** of Σ.

We prove that Δ has the **finite intersection property**. Both parts of Δ are closed under intersection. We show that $V(\phi) \cap Y \neq \emptyset$. Since $\phi \in \mathsf{con}(\Sigma)$, there is v with $R^{\mathsf{ue}}uv$ and $\mathfrak{ue}(\mathfrak{M}), v \models \phi$ by assumption. $pre^d(Y) \in u$ implies $Y \in v$ (Lemma 5.52). So we have that $V(\phi) \cap Y \neq \emptyset$.

Again, by the Ultrafilter Theorem 5.46 there is an ultrafilter u' with $\Delta \subseteq u'$ where u' satisfies all required properties. □

The main theorem tells us that modal equivalence of models can **completely be characterized** by bisimulations in their ultrafilter extensions.

Theorem 5.56 (Bisimulation-somewhere-else)

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be two models of the same similarity type, $w \in W_{\mathfrak{M}}$ and $w' \in W_{\mathfrak{M}'}$. Then, we have the following:

$$\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w' \quad \text{iff} \quad \mathfrak{ue}(\mathfrak{M}), \pi_w \leftrightarrows \mathfrak{ue}(\mathfrak{M}'), \pi_{w'}.$$

Proof.

Result of Proposition 5.55, Proposition 5.53, and Theorem 5.36:

$\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w'$ iff $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{ue}(\mathfrak{M}'), \pi_w$ and
$\mathfrak{M}', w' \leftrightsquigarrow \mathfrak{ue}(\mathfrak{M}), \pi_{w'}$ iff $\mathfrak{ue}(\mathfrak{M}), \pi_w \leftrightsquigarrow \mathfrak{ue}(\mathfrak{M}'), \pi_{w'}$
iff $\mathfrak{ue}(\mathfrak{M}), \pi_w \leftrightarrows \mathfrak{ue}(\mathfrak{M}'), \pi_{w'}$ □

5.3 Saturated Models and Ultraproducts

Basics of FOL Model Theory

Our main result relates the **modal language** to **first-order language** therefore we need some basic **model theoretic concepts** about FOL. Also the **standard translation** (see Slide 183 ff) is needed.

Recall that a **FOL model** was given by (U, I) where U is the **universe** or **domain** and I the **interpretation function**. From now on, we usually use $\mathcal{A}$ and $\mathcal{A}'$ to denote FOL models.

Let $\mathcal{A} = (U, I)$ and $\mathcal{A}' = (U', I')$ be two FOL models for the same first-order language. Both models are said to be **isomorphic** if there is a **bijective** function $f : U \to U'$ which satisfies the following properties:

1. For each **predicate symbol** P (k-ary) it holds that
$$I(P)(u_1, \ldots, u_k) \quad \text{iff } I'(P)(f(u_1), \ldots, f(u_k))$$
2. For each **function symbol** F (k-ary) it holds that
$$f(I(F)(u_1, \ldots, u_k)) = I'(F)(f(u_1), \ldots, f(u_k))$$
3. For each **variable** it holds that
$$f(I(x)) = I'(x)$$

f is called an **isomorphism** between $\mathcal{A}$ and $\mathcal{A}'$, short: $f : \mathcal{A} \to \mathcal{A}'$.

Inuitively: Isomorphic models can be considered to be “the same”.

We use the notation $\varphi(x_1, \ldots, x_n)$ to indicate that φ includes the **free variables** $x_1, \ldots, x_n$. Moreover, given n constants $u_1, \ldots, u_n$ we use

$$\mathcal{A} \models \varphi(x_1, \ldots, x_n)[u_1, \ldots, u_n]$$

as shortcut for $\mathcal{A}^{[x_1/u_1]\ldots[x_n/u_n]} \models \varphi$. Often we also just write

$$\mathcal{A} \models \varphi[u_1, \ldots, u_n]$$

for $\mathcal{A} \models \varphi(x_1, \ldots, x_n)[u_1, \ldots, u_n]$ if the variables of φ are clear from context.

Proposition 5.57

Let f be an **isomorphism** *between $\mathcal{A}$ and $\mathcal{A}'$. Then for all $\varphi(x_1, \ldots, x_n)$ and $u_1, \ldots, u_n \in U_{\mathcal{A}}$ we have that*

$$\mathcal{A} \models \varphi[u_1, \ldots, u_n] \quad \textit{iff} \quad \mathcal{A}' \models \varphi[f(u_1), \ldots, f(u_n)]$$

Definition 5.58 (Elementarily Equivalent)

Let $\mathcal{L}$ be a FOL language. Two models are said to be $\mathcal{L}$-elementarily equivalent if they satisfy the same formulae of $\mathcal{L}$.

Does elementary equivalence imply isomorphism? What if the models are finite?

Proposition 5.59 (Elementary Eq. and Isom.)

If $\mathcal{A}_1 \cong \mathcal{A}_2$ then $\mathcal{A}_1, \mathcal{A}_2$ are $\mathcal{L}$-**elementary equivalent**. *If the case of finite structures the converse does also hold.*

Analogously to Definition 5.8 we define a first-order submodel. The functions and predicates have just to be restricted to the respective subset of the universe. Whenever $\mathcal{A}$ is a submodel of $\mathcal{A}'$ we call $\mathcal{A}'$ an extension of $\mathcal{A}$.

Definition 5.60 (Elementary Ext., -Embedding)

We call a model $\mathcal{A}'$ an elementary extension of $\mathcal{A}$ ($\mathcal{A} \preceq \mathcal{A}'$) if the following holds:

1. $\mathcal{A}'$ is an **extension** of $\mathcal{A}$, and
2. for any formula $\varphi(x_1, \dots, x_n)$ and elements $u_1, \dots, u_n \in \mathcal{A}$ the following holds:

$$\mathcal{A} \models \varphi[u_1, \dots, u_n] \quad \text{iff} \quad \mathcal{A}' \models \varphi[u_1, \dots, u_n]$$

We call a function $f : U \to U'$ an elementary embedding of $\mathcal{A}$ to $\mathcal{A}'$ if for any formula $\varphi(x_1, \dots, x_n)$ and elements $u_1, \dots, u_n \in \mathcal{A}$ it holds that

$$\mathcal{A} \models \varphi[u_1, \dots, u_n] \quad \text{iff} \quad \mathcal{A}' \models \varphi[f(u_1), \dots, f(u_n)]$$

Next, we extend a given language by **new** constants which allow to refer to specific elements of the model.

Definition 5.61 ($\mathcal{L}[A]$, Expansion)

Let $\mathcal{A} = (U, I)$ be a first-order model for $\mathcal{L} \subseteq \mathcal{L}_{FOL}$ and $A \subseteq U$ (so, A is a subset of the universe).

Then $\mathcal{L}[A]$ denotes the language which is defined by $\mathcal{L}$ **extended** by the **new** constants c_a, one for each $a \in A$.

The *expansion of $\mathcal{A}$ with respect to A* (denoted by $\mathcal{A}_A$) extends $\mathcal{A}$ by the new constants c_a which are interpeted as a for all $a \in A$ (i.e. $\mathcal{A}_A(c_a) = a$).

That is, the expansion includes new constants which allow to refer to specific individuals. *All elements have names and can be used in formulae.*

The next proposition characterizes elementary extensions: Models which are elementarily equivalent must be "very similar".

Proposition 5.62

$\mathcal{A}'$ is an **elementary extension** *of $\mathcal{A} = (U, I)$ iff $\mathcal{A}$ is a* **submodel** *of $\mathcal{A}'$ and $\mathcal{A}_U$ is* **isomorphic** *to $\mathcal{A}'_U$.*

Theorem 5.63

Let $\mathcal{A}$ be a model and let $\#(\mathcal{L}) \leq \mathfrak{c} \leq |U|$ where $\#(\mathcal{L})$ denotes the number of **non-logical symbols** *in $\mathcal{L}$. Then, there is an* **elementary submodel** *of cardinality $\mathfrak{c}$. Furthermore, for $X \subseteq U$, $|X| \leq \mathfrak{c}$ there is an elementary submodel of $\mathcal{A}$ with cardinality $\leq \mathfrak{c}$ which contains X.*

Saturated Models

Definition 5.64 ($\mathcal{L}^1_{FOL}$, free, x-type)

We define $\mathcal{L}^1_{FOL} \subset \mathcal{L}_{FOL}$ as the first-order sublanguage in which each formula has exactly **one free variable**. (A variable is called **free** if it is not in the scope of any quantifier).

Let $\Sigma(x) \subseteq \mathcal{L}^1_{FOL}$ is called **x-type** if x is the only free variable in $\Sigma(x)$. We will often just use **type** to refer to an x-type.

Definition 5.65 ((Finite) Realization)

Let $\Sigma(x)$ be an **x-type** and let $\mathcal{A} = (U, I)$ be a first-order model.
We say that $\mathcal{A}$ **realizes** $\Sigma(x)$ if **there is** a $u \in U$ so that **for all** $\varphi \in \Sigma(x)$ it holds that $\mathcal{A} \models \varphi[x]$. $\Sigma(x)$ is said to be **finitely realizable** in $\mathcal{A}$ if $\mathcal{A}$ realizes every **finite** subset of Σ.

The next defintion characterizes how "**expressive**" a FOL model is. Suppose we can refer to any element of the universe of the model. Can the model **realize** any reasonable set of formulae which we can build?

Definition 5.66 (Countably Saturated)

A model $\mathcal{A}$ is called **countably saturated** if **for every** subset $A \subseteq W_{\mathcal{A}}$ the **expansion** $\mathcal{A}_A$ **realizes** every x-type $\Gamma \subseteq \mathcal{L}^1_{FOL}[A]$ that is **consistent** with the first-order theory of $\mathcal{A}_A$.

Note: There also is the notion of α-saturation where one requires that $|A| \leq \alpha$. However, for our needs countable saturation suffices.

Example 5.67

Every finite model $\mathcal{A}$ is countably saturated.

Let $\Gamma(x)$ be an x-type consistent with the theory of $\mathcal{A}$. That is, there is a model $\mathcal{A}'$ with $\mathcal{A}' \models Th(\mathcal{A}) \cup \Sigma(x)$.
Then, there is an **elementary extension** of $\mathcal{A}'$ for $Th(\mathcal{A}) \cup \Gamma(x)$. Hence, $\mathcal{A}'$ **realizes** $\Gamma(x)$ but then also $\mathcal{A}$ as both models must be isomorphic (cf. Prop. 5.59).

Example 5.68 ($(\mathbb{Q}, <)$ is countably saturated)

The model $\mathcal{A} = (\mathbb{Q}, <)$ is countably saturated.

Consider the FOL language $\mathcal{L}$ over $<$ and $=$. Let $A \subseteq \mathbb{Q}$ and $\Sigma(x)$ be an x-type of $\mathcal{L}[A]$ **consistent** with $Th(\mathcal{A}_A)$. Then, there is a model $\mathcal{A}'$ of $Th(\mathcal{A}_A)$ which **realizes** $\Sigma(x)$. Now, let $\mathcal{A}''$ be a countably **elementary** submodel of $\mathcal{A}'$ that **realizes** $\Sigma(x)$.

How does $\mathcal{A}''$ look like? The model has to be a **countable dense linear ordering without endpoints**! (Why?). Hence, $\mathcal{A}''$ and $\mathcal{A}$ are **isomorphic**! Hence, $\mathcal{A}_A$ realizes $\Sigma(x)$.

So, can we give a similar proof that $(\mathbb{N}, <)$ is countably saturated?

Example 5.69 ($(\mathbb{N}, <)$ is not countably saturated)

Consider the model $\mathcal{A} = (\mathbb{N}, <)$ and define the following set of formulae

$$\Gamma(x) := \{\exists y_1(y_1 < x), \ldots, \exists y_1 \ldots y_n(y_1 < \cdots < y_n < x), \ldots\}$$

It is easy to see that $\Gamma(x)$ and $\mathcal{A}$ are consistent as $\Sigma(x)$ is finitely realizable in $\mathcal{A}$ (see Proposition 5.70). (Can you come up with a model?) However, $\Gamma(x)$ is obviously not realizable in $\mathcal{A}$.

Proposition 5.70 (Consistency & Realization)

Let $\mathcal{A}$ be a FOL-model and Σ be an x-type of $\mathcal{L}^1[A]$ where $A \subseteq U$. Then, the following holds:
*Σ is consistent with $Th(\mathcal{A}_A)$ **iff** Σ is finitely realizable in $\mathcal{A}_A$.*

Proof.

Σ is **consistent** with $Th(\mathcal{A}_A)$

iff $Th(\mathcal{A}_A) \cup \Sigma \not\models \bot$

iff $\exists \mathcal{A}' \ (\mathcal{A}' \models Th(\mathcal{A}_A) \cup \Sigma(x))$

iff $\exists \mathcal{A}' \ \forall \Sigma' \subseteq \Sigma, |\Sigma'| < \infty (\mathcal{A}' \models Th(\mathcal{A}_A) \cup \{\exists x \bigwedge \Sigma'(x)\})$

iff $\forall \Sigma' \subseteq \Sigma, |\Sigma'| < \infty (\mathcal{A}_A \models Th(\mathcal{A}_A) \cup \{\exists x \bigwedge \Sigma'(x)\})$

iff Σ is **finitely realizable** in $\mathcal{A}_A$ □

Ultraproducts and Ultrapowers

Let I be a non-empty set. Moreover, let for each $i \in I$, W_i be a non-empty set as well. We define the following **Cartesian product**:

$$C := \prod_{i \in I} W_i$$

We can also interpret C as the **set of functions** $f : I \to \bigcup_{i \in I} W_i$ such that $f(i) \in W_i$. We also write $f \in C$ to refer to such a function. In the following we assume that C is always defined as above.

Definition 5.71 (U-Equivalence)

Let U be an **ultrafilter** over I and $f, g \in C$. Functions f and g are called **U-equivalent**, $f \sim_U g$, if $\{i \in I \mid f(i) = g(i)\} \in U$.

Note: $\sim_U$ is an equivalence relation.

Definition 5.72 (Ultraproduct of Sets)

Let U be an ultrafilter over I. The **ultraproduct $\prod_U W_i$ of W_i modulo U** is defined as follows:

$$\prod_U W_i := \{[f]_U \mid f \in \prod_{i \in I} W_i\}$$

If there is a W such that $W_i = W$ for all $i \in I$ the ultraproduct is called **ultrapower** and is denoted by $\prod_U W$.

Note: $[f]_U$ denotes the **equivalence class** of f regarding $\sim_U$.
We now transfer the construction to models.

Definition 5.73 (Ultraproduct of Models)

Let $\mathfrak{M}_i = (W_i, R_i, V_i)$ be models for $i \in I$ and U be an ultrafilter over I. The **ultraproduct $\prod_U \mathfrak{M}_i$ of $\mathfrak{M}_i$ modulo U** is the following model (W_U, R_U, V_U):

1. $W_U = \prod_U W_i$ (W_U is the ultraproduct of the worlds)
2. V_U is defined as follows:

$$[f]_U \in V_U(\mathsf{p}) \quad \text{iff} \quad \{i \in I \mid f(i) \in V_i(\mathsf{p})\} \in U$$

3. R_U is defined as follows:

$$R_U[f]_U[f']_U \quad \text{iff} \quad \{i \in I \mid R_i f(i) f'(i)\} \in U$$

So, where do we need ultraproducts for?

The next proposition states that a model is **modally equivalent** to its ultrapower for any ultrafilter over any non-empty index set I.

Proposition 5.74 (Models and Ultrapowers)

Let U be an ultrafiler over a non-empty I, $\mathfrak{M}$ be a model and $\prod_U \mathfrak{M}$ its **ultrapower modulo U**. *Then,*

$$\mathfrak{M}, w \text{ and } \prod_U \mathfrak{M}, [f_w]_U \text{ are modally equivalent}$$

where f_w denotes the **constant function** $i \mapsto w$ *for all* $i \in I$.

In the next step, we show that the ultrapower is also **countably saturated** (interpreted as FOL model) if we use a **particular** ultrafilter over I.

Proof.

The proof is done by structural induction. We consider the two cases, the remaining are left as exercise.

$$\begin{aligned}
\mathfrak{M}, w \models \mathsf{p} \text{ iff } & w \in V(\mathsf{p}) \\
\text{iff } & \{i \in I \mid f_w(i) \in V(\mathsf{p})\} = I \\
\text{iff } & \{i \in I \mid f_w(i) \in V(\mathsf{p})\} = I \in U \\
\text{iff } & [f]_U \in V_U(\mathsf{p}) \\
\text{iff } & \prod_U \mathfrak{M}, [f_w]_U \models \mathsf{p}
\end{aligned}$$

□

Let us consider the case $\varphi = \Diamond\, \psi$

"$\Rightarrow$": Let $\mathfrak{M}, w \models \Diamond\, \psi$. Then, there is w' with Rww' and $\mathfrak{M}, w' \models \psi$. So, we have $\{i \in I \mid R_i f_w(i) f_{w'}(i)\} = I$ which gives us that $R_U[f_w][w_{w'}]$. By induction we have $\prod_U \mathfrak{M}, [f_{w'}] \models \psi$ which proves that $\prod_U \mathfrak{M}, [f_w] \models \Diamond\, \psi$.

"$\Leftarrow$": Let $\prod_U \mathfrak{M}, [f_w] \models \Diamond\, \psi$. Then, there is a w' such that $R_U[f_w][f_{w'}]$ and $\prod_U \mathfrak{M}, [f_{w'}] \models \psi$ which implies that $\emptyset \neq \{i \in I \mid R_i ww'\} \in U$ hence Rww' and by induction $\mathfrak{M}, w' \models \psi$ and therewith $\mathfrak{M}, w \models \Diamond\, \psi$.

Definition 5.75 (Countable Incompleteness)

An **ultrafilter** is called *countably incomplete* if it is **not** closed under **countable** intersections.

Note: It still must be closed under **finite** intersections otherwise it would not be an ultrafilter!

Example 5.76 (Countably Incomplete Uf)

Let U be an ultrafilter over $\mathbb{N}$ which does not contain any singleton $\{n\}$ for $n \in \mathbb{N}$. (Such an ultrafilter exists! $\rightsquigarrow$ Exercise). As U is an ultrafiler we have that $\mathbb{N}\backslash\{n\} \in U$ for all $n \in \mathbb{N}$ but

$$\emptyset = \bigcap_{n\in\mathbb{N}} (\mathbb{N}\backslash\{n\}) \notin U$$

otherwise U would be proper and therefore $\{n\} \in U$.

We state the following theorem without proof. The construction of an ultraproduct for **first-order logic** is done in a similar way as for modal logic.

Lemma 5.77

Let $\mathcal{L}$ be a **countable** *first-order language, U be a* **countably incomplete** *ultrafilter over $I \neq \emptyset$, and $\mathcal{A}$ be a FOL model for $\mathcal{L}$. Then, the ultrapower $\prod_U \mathcal{A}$ is* countably saturated.

Proof.

See Chang/Keisler: Model Theory. □

5.4 Characterization Theorem

In this section, we prove the following result:

> *Modal logic is the* bisimulation-invariant *fragment of first-order logic.*

For our main result, we need some abstract concepts which were introduced in the previous section:

- Basics of first-order model theory,
- Saturated models, and
- Ultraproducts.

The following result states that **modal saturation** can be seen as modal counterpart of **countable saturation** in FOL.

Theorem 5.78

Let $\mathfrak{M}$ be τ-model. If the first-order interpretation of $\mathfrak{M}$ is **countably saturated** *then $\mathfrak{M}$ is* **modally saturated**. *Hence, the class of such models has the* Hennessy-Milner property.

Proof.

We tackle the basic modal logic case. Let $\mathcal{A}$ be **countably saturated** and be the first-order representation of $\mathfrak{M} = (W, \mathcal{R}, V)$. Let $w \in W$ and $\Sigma \subseteq \mathcal{L}_{BML}$ be a set of formulae **finitely satisfiable** in $\{w' \mid Rww'\}$ of $\mathfrak{M}$. Let

$$\Sigma' := \{Rc_a x\} \cup \{ST_x(\varphi) \mid \varphi \in \Sigma\}$$

- Σ' is **consistent** with $Th(\mathcal{A}_a)$ as Σ' is finitely realizable.
- Hence, Σ' **realizable** in some $b \in \mathcal{A}$.
- Since, in particular, $\mathcal{A} \models Rc_a b$, b must be a **successor of** a.
- By Theorem **??** and $\mathcal{A} \models ST_x(\varphi)[b]$ for all $\phi \in \Sigma$ we have $\mathfrak{M}, b \models \Sigma$.

This shows that $\mathfrak{M}$ is modally saturated! □

In Theorem 5.56 we have shown that models are modally equivalent iff there is a bisimulation **somewhere-else**, namely in the **ultrafilter extension**. Here we show that somewhere-else can also mean in the **ultrapower**.

Theorem 5.79 (Bisimulation in Ultrapowers)

Let $\mathfrak{M}$ and $\mathfrak{M}'$ be modal models of the same similarity type, $w \in \mathfrak{M}$ and $w' \in \mathfrak{M}'$. Then, the following are equivalent:

1. $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w'$
2. *There are* **ultrapowers** *(i.e. an appropriate ultrafilter U over a non-empty set I) $\prod_U \mathfrak{M}$ and $\prod_U \mathfrak{M}'$ such that $\prod_U \mathfrak{M}, [f_w]_U \leftrightarrows \prod_U \mathfrak{M}', [f_{w'}]_U$ where f_w denotes the constant function $i \mapsto w$ for all $i \in I$ and $f_{w'}$ is defined analogously.*

Proof.

"(2)⇒(1)": By Proposition 5.74 we have:

$$\mathfrak{M}, w \leftrightsquigarrow \prod_U \mathfrak{M}, [f_w]_U \quad \text{and} \quad \mathfrak{M}', w' \leftrightsquigarrow \prod_U \mathfrak{M}', [f'_w]_U$$

and since bisimulation implies modal equivalence we also have that $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w'$ as $\leftrightsquigarrow$ is transitive.

□

“(1)⇒(2)”: Assume $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', w'$. We construct bisimular ultrapowers. Firstly, take $I := \mathbb{N}$ as **index set** and let U be a **countably incomplete** ultrafilter over I (cf. Example 5.76). By Lemma 5.77 the ultrapowers $\prod_U \mathfrak{M}$ and $\prod_U \mathfrak{M}'$ are countably saturated.

By Proposition 5.74 and the assumption that $w \leftrightsquigarrow w'$ we have that $\prod_U \mathfrak{M}, [f_w]_U \leftrightsquigarrow \prod_U \mathfrak{M}', [f_{w'}]_U$.

Finally, as $\prod_U \mathfrak{M}$ and $\prod_U \mathfrak{M}'$ are countably saturated by Theorem 5.78 it follows that $\prod_U \mathfrak{M}$ and $\prod_U \mathfrak{M}'$ are **modally saturated**. Hence, they have the Hennesey-Millner property and thus are bisimular.

Characterization Theorem

We know the following results:

1. FOL is **equivalent** to the hybrid language with renaming
2. The modal language is **weaker** than FOL.

The characterization theorem exactly identifies which **fragment** of FOL is **equally expressive** as the modal language!

In order to prove this theorem we need all the theoretical tools presented in this and the previous section: This indicates the depth of this result.

Let $\mathfrak{M}$ be a **modal model** then we use $\mathfrak{M}_F$ to denote its *first-order representation*. Analogously, we use $\mathcal{A}_M$ to denote the *modal representation* of a **first-order model** $\mathcal{A}$.

The following Lemma gives another characterization of modal equivalence. The characterization in terms of **elementary embeddings** is important because satisfaction of FOL formulae is **not closed under ultrafilter extensions** but it is **under elementary embeddings**!

Lemma 5.80 (Detour Lemma)

Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be ***modal models****, $w_1 \in W_1$ and $w_2 \in W_2$. Then, the following statements are all equivalent:*

1. $\mathfrak{M}_1, w_1 \leftrightsquigarrow \mathfrak{M}_2, w_2$
2. $\mathfrak{ue}(\mathfrak{M}_1), \pi_{w_1} \leftrightarrows \mathfrak{ue}(\mathfrak{M}_2), \pi_{w_2}$
3. *There are* ***countably saturated*** *models $\mathcal{A}_1'$ and $\mathcal{A}_2'$, states $w_1' \in U_1'$, $w_2' \in U_2'$ and* ***elementary embeddings*** $f_1 : (\mathfrak{M}_1)_F \preceq \mathcal{A}_1'$ *and* $f_2 : (\mathfrak{M}_2)_F \preceq \mathcal{A}_2'$ *such that*
   1. $f_1(w_1) = w_1'$ *and* $f_2(w_2) = w_2'$*, and*
   2. $(\mathcal{A}_1')_M, w_1' \leftrightarrows (\mathcal{A}_2')_M, w_2'$

Proof.

"(1)⇔(2)": Bisimulation somewhere-else Theorem 5.56

"(1)⇔(3)": By Theorem 5.79, (1) iff there are **ultrapowers** $\prod_U \mathfrak{M}$ and $\prod_U \mathfrak{M}'$ such that $\prod_U \mathfrak{M}, [g_w]_U \leftrightarrows \prod_U \mathfrak{M}', [g_{w'}]_U$ where g_w denotes the *constant function* $i \mapsto w$ for all $i \in I$.

We just chose $\mathfrak{M}'_1$ and $\mathfrak{M}'_2$ as the ultraproducts of $\mathfrak{M}$ and $\mathfrak{M}'$ and define $f_1(w_1) := [g_{w_1}]_U$ and $f_2(w_2) := [g_{w_2}]_U$. The functions are elementary embeddings and the ultrapowers can also be constructed as countably saturated models (proof of Theorem 5.79).

□

Now, we are ready to prove our main theorem:

> *Modal logic is the* bisimulation-invariant *fragment of the first-order language.*

Formally, the next definition states how **bisimulation-invariance** is understood for first-order logic.

Definition 5.81 (Invariant for Bisimulation)

Let $\{\varphi(x)\}$ be an x-type. $\varphi(x)$ is invariant for bisimulation if for all **modal models** $\mathfrak{M}$ and $\mathfrak{M}'$, all worlds $w \in \mathfrak{M}$ and $w' \in \mathfrak{M}'$, and all bisimulations $\mathcal{B} \subseteq \mathfrak{M} \times \mathfrak{M}'$ with $w\mathcal{B}w'$ it holds that

$$\mathfrak{M}_F \models \varphi(x)[w] \text{ iff } \mathfrak{M}'_F \models \varphi(x)[w'].$$

The final theorem of this section is named after the logician *Johan van Benthem*.

Theorem 5.82 (Characterization Theorem)

Let $\{\varphi(x)\}$ be an x-type. Then, the following are equivalent:

1. $\varphi(x)$ *is* invariant for bisimulation.
2. $\varphi(x)$ *is* equivalent *to the standard translation of a modal formula.*

Proof.

The easy part first. "$\Leftarrow$": Let φ be the standard translation of a modal formula. Clearly, by Theorem 5.26 we have that φ is invariant under bisimulation. □

"$\Rightarrow$": Let $\alpha(x)$ be a **FOL formula invariant for bisimulation**. Consider the **modal consequences** of α:

$$Cn(\alpha) := \{ST_x(\varphi) \mid \varphi \in \mathcal{L}_{\mathsf{ML}} \text{ and } \alpha(x) \models ST_x(\varphi)\}$$

- Claim 1: If $Cn(\alpha) \models \alpha(x)$ then $\exists \varphi \in \mathcal{L}_{\mathsf{ML}}$ such that $\alpha(x)$ is equivalent to $ST_x(\varphi)$. Proof: Assume that $Cn(\alpha) \models \alpha(x)$ then by the **compactness theorem** there is a finite $X \subseteq Cn(\alpha)$ such that $X \models \alpha(x)$ and hence $\models \bigwedge X \rightarrow \alpha(x)$ and trivially also the reverse implication thus $\bigwedge X \leftrightarrow \alpha(x)$. Finally, it is easy to see that $\exists \varphi \in \mathcal{L}_{\mathsf{ML}}$ with $\bigwedge X = ST_x(\varphi)$.

So, it remains to show that $Cn(\alpha) \models \alpha(x)$. Proof: Assume $\mathcal{A} \models Cn(\alpha)[u]$ with $u \in \mathcal{A}$. We have to show that $\mathcal{A} \models \alpha(x)[w]$. We set

$$T(x) := \{ST_x(\varphi) \mid \mathcal{A} \models ST_x(\varphi)[w]\}$$

- Claim 2.1: $T(x) \cup \{\alpha(x)\}$ is consistent. Proof: Assume the opposite. By the compactness theorem there is a finite $T' \subseteq T(x)$ such that $\models \alpha(x) \rightarrow \neg \bigwedge T'(x)$, hence, $\neg \bigwedge T'(x) \in Cn(\alpha)$. This implies $\mathcal{A} \models \neg \bigwedge T'(x)[w]$. Contradiction!

So, let $\mathcal{B} \models T(x) \cup \{\alpha(x)\}[v]$. Now, it is easy to see that $\mathcal{A}_M, w$ and $\mathcal{B}_M, v$ are **modally equivalent**. If modal equivalence implied bisimularity we would be done! (But this is not the case!)

However, by the Detour Lemma, we have that there are two **countably saturated** models $\mathcal{A}'$ and $\mathcal{B}'$ such that $\mathcal{A}$ and $\mathcal{B}$ are **elementarily embedded**, respectively, and states $w' \in \mathcal{A}'$ and $v' \in \mathcal{B}'$ such that $\mathcal{A}'_M, w' \leftrightarrows \mathcal{B}'_M, v'$. Now, satisfaction of FOL formulae is preserved under **elementary embeddings**. (Which is not the case for ultrafilter extension! And that is the reason we need ultraproducts!)

Hence, $\mathcal{B} \models \alpha(x)[v]$ implies that $\mathcal{B}' \models \alpha(x)[v']$ and hence $\mathcal{A}' \models \alpha(x)[w']$ because $\alpha(x)$ is **invariant for bisimulation** by assumption. Now we finally have that $\mathcal{A} \models \alpha(x)[w]$.

We end this section with a small summary of the basic relationships.

Summary

- **Bisimulation** implies modal equivalence.
- In general, **modal equivalence** does **not** imply bisimulation.
- **Ultrafilter extensions completely** characterize modal equivalence.
- However, they are **too weak for the Characterization Theorem** as satisfaction of FOL formulae is **not** closed under ultrafilter extensions.
- The **Detour Lemma** also completely characterizes modal equivalence but in terms of **elementary embeddings** (which preserve satisfaction of FOL formulae).
- **Ultraproducts** were needed to **prove the existence** of these elementary embeddings.