TECHNISCHE UNIVERSITÄT WIEN

DIPLOMARBEIT

Interpolation durch Übersetzung

zur Erlangung des akademischen Grades

Diplom-Ingenieur/in

im Rahmen des Studiums

Technische Mathematik

eingereicht von

Charlotte Jergitsch
Matrikelnummer 11779508

ausgeführt am Institut für Diskrete Mathematik und Geometrie
der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuung: Ao.Univ.Prof. Dr.phil. Matthias Baaz

Wien, 20.03.2023

(Unterschrift Verfasser/in) (Unterschrift Betreuer/in)

MASTER'S THESIS

Interpolation by Translation

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur/in

in

Technische Mathematik

by

Charlotte Jergitsch
Registration Number 11779508

to the Institute of Discrete Mathematics and Geometry
at the Faculty of Mathematics and Geoinformation of the Vienna University of Technology

Advisor: Ao.Univ.Prof. Dr.phil. Matthias Baaz

Vienna, 20.03.2023

(Signature of Author) (Signature of Advisor)

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Arbeit überein.

Wien, 20.03.2023

Unterschrift

Acknowledgements

I want to thank my supervisor Matthias Baaz for providing his guidance and his broad knowledge, be it in person or via telephone from all kinds of places on this globe.

I am also deeply grateful to my parents, Beatriz and Friedrich, for their love and relentless support. You have my back at all times and this means the world to me.

Further I want to thank my family and friends who have all been such a source of encouragement and joy. With you everything is possible.

Abstract

Craig's Interpolation Theorem is a fundamental result in classical first-order logic. In this master's thesis we examine interpolation in the quantified modal logic $\mathcal{S}5$. Craig's Interpolation Theorem does not generally hold in this logic, but we will examine some exceptions.

First, we describe Gentzen's sequent calculus and important results such as the cut-elimination theorem and Craig's Interpolation Theorem. We explore the connection to Beth's Definability Theorem, which is implied by the Interpolation Theorem. We proceed to introduce the modal logic $\mathcal{S}5$ and explore some possibilities and limitations of different calculi. Then we present Kit Fine's proof of the statement that the Interpolation Theorem does not hold in $\mathcal{S}5$. Fine proceeds by showing that Beth's Definability Theorem is not valid in $\mathcal{S}5$. However, there are exceptions. In the last chapter, we show that one can find the interpolant for fragments of quantified $\mathcal{S}5$, namely the fragment of sequents of prenex formulas and those containing weak modal operators only. We find the interpolants for these sequents by translating them to classical two-sorted first-order logic, finding the interpolant there with Craig's Interpolation Theorem, and then translating the interpolant back to $\mathcal{S}5$.

Kurzfassung

Craigs Interpolationstheorem ist ein zentrales Resultat in der klassischen Prädikatenlogik erster Stufe. In dieser Diplomarbeit betrachten wir die Möglichkeiten der Interpolation in der quantifizierten Modallogik $\mathcal{S}5$. Craigs Interpolationstheorem gilt in dieser Logik nicht im Allgemeinen und wir werden uns in dieser Arbeit mit einigen Ausnahmen befassen.

Zuerst beschreiben wir Gentzens Sequenzialkalkül und wichtige darauf beruhende Resultate, wie Gentzens Hauptsatz über die Schnittelimination sowie Craigs Interpolationstheorem. Wir betrachten die Verbindung zum Beth'schen Definierbarkeitstheorem, welches aus dem Interpolationstheorem folgt. Im darauf folgenden Kapitel führen wir die Modallogik $\mathcal{S}5$ ein und betrachten die Möglichkeiten und Limitationen verschiedener Kalküle. Anschließend präsentieren wir den Beweis von Kit Fine, dass das Interpolationstheorem für die quantifizierte Modallogik $\mathcal{S}5$ nicht gilt. Der Beweis gelingt Fine, indem er zeigt, dass das Beth'sche Theorem in $\mathcal{S}5$ nicht zutrifft. Dazu gibt es jedoch Ausnahmen. Im letzen Kapitel zeigen wir, dass man für Sequenzen aus dem Fragment der pränexen Formeln sowie aus dem Fragment der Formeln mit ausschließlich schwachen Modaloperatoren sehr wohl Interpolanten finden kann. Wir finden diese Interpolanten, indem wir die Sequenzen in die klassische zweisortige Logik übersetzen, in dieser mithilfe von Craigs Interpolationstheorem die Interpolante finden, und diese schließlich in $\mathcal{S}5$ zurückübersetzen.

Contents

1 Introduction

Craig's Interpolation Theorem is a fundamental result in classical first-order logic. The interpolant of an implication shows the essence of the logical connection of the premise and the consequence. Further, the Interpolation Theorem implies Beth's Definability Theorem, another cornerstone in the formalization of mathematics. Thus, being able to show that the Interpolation Theorem holds in a logic is of great interest. However, it does not apply to some logics because they do not allow for calculi that provide the necessary proof structure to construct an interpolant from. One of them is the quantified modal logic $\mathcal{S}5$, which we will examine further in this thesis.

In the early 19th century, a movement to rigorize mathematics started, pushed by mathematicians such as Cauchy and Bolzano. This movement culminated in the development of the formal axiomatic method for mathematical reasoning. ([Swi98], p. 760). In the process of formalizing logical deductions, Gentzen introduced his sequent calculus in his paper "Untersuchungen über das logische Schließen" in 1935 ([Gen35]). Before that, Hilbert's deduction system, which was developed with contributions from Frege, Russell and Hilbert himself, was known and used as a way to formalize logical deductions. These two systems differ in the sense that Hilbert-style deduction systems allow a large number of logical axiom schemes as a base but only few inference rules to deduce proofs from these axioms (often just the modus ponens and the generalization rule for predicate logic). Gentzen felt dissatisfied with this style of deduction, as he thought that it does not correspond to the way humans normally go about conducting mathematical proofs. He developed a formalism that he deemed more "natural", the "calculus of natural deduction" ([Gen35] p. 176). In contrast to Hilbert's system, the natural deduction system uses only very few axiom schemes but a greater number of inference rules for its proofs. Whereas the Hilbert-style systems work with unconditional tautologies as starting points, natural deduction uses hypotheses as a starting points, i.e. conditional tautologies. Exploring the properties of the natural deduction system led Gentzen to formulate his famous "Hauptsatz", also known today as the cut-elimination theorem (Theorem (2.7)). However, the natural calculus turned out not to be suitable to prove the Hauptsatz, which led Gentzen to introduce the sequent calculus. Gentzen was able to formulate and prove the Hauptsatz for this new calculus. It is equivalent to his natural deduction calculus, but thanks to the Hauptsatz has some convenient properties that the latter does not have. The Hauptsatz is remarkable as it expresses that every purely logical proof can be transformed into a normal form such that the proof does not contain any detours. Consequently, all expressions used in the proof are part of the final result, such that the proof contains only expressions that necessarily need to be introduced. Another benefit of this property is that as a consequence, the length of proofs is bounded in the sequent calculus. Every inference step introduces a new logical symbol that will be contained in the endresult, such that the length of the proof is limited by the structure of the final formula in the deduction.

We refer to such calculi that allow for cut elimination and thus proofs without detours as "analytic". The property of analyticity has far-reaching consequences as Gentzen's Hauptsatz lays the foundation for several other important theorems, foremost Craig's Interpolation Theorem for classical first-order logic. This theorem tells us that if we have a valid implication $\varphi \supset \psi$, where both arguments of the implication have at least one predicate symbol in common, then there exists a formula γ in the common language of the arguments, called an "interpolant", such that both formulas $\varphi \supset \gamma$ and $\gamma \supset \psi$ are valid. Thus, the interpolant reveals the logical connection of φ and ψ that makes up the implication. Craig derived his interpolation theorem originally in order to prove Beth's Definability Theorem, which immediately follows from Craig's theorem. Beth's Definability Theorem states that a concept is implicitly definable in a given logic if and only if it is explicitly definable. This is a central result about the definability of concepts and being implied by Craig's Interpolation Theorem, the latter becomes particularly interesting for other logics as well.

Craig's Interpolation Theorem does not hold for all logics however. For example, it is not true for the quantified modal system $\mathcal{S5}$. In general, there is no analytic calculus known for $\mathcal{S5}$, meaning that we cannot eliminate the cut-rule in a sequent calculus for $\mathcal{S5}$ without reducing its expressive power. Consequently, we would not be able to derive the statement of Craig's Interpolation Theorem for $\mathcal{S5}$ in the same manner as in classical first-order logic. In 1979, Kit Fine went further and showed in his paper "Failures of the Interpolation Lemma in quantified Modal Logic" that Craig's Interpolation Theorem does not hold in quantified $\mathcal{S5}$ by showing that Beth's Definability Theorem fails for it ([Fin79]). However, there are some exceptions – there are certain classes of formulas in $\mathcal{S5}$ to which Craig's Interpolation Theorem applies. In this thesis, we will show that amongst others, prenex formulas in $\mathcal{S5}$ do interpolate. We will deduce this result by translating the formulas from $\mathcal{S5}$ to two-sorted classical first-order logic, finding the interpolant there and translating it back to $\mathcal{S5}$. Due to the prenex structure of the formula, which gets preserved by the translation, it is possible to find an interpolant in prenex form in first-order logic that can be translated back to $\mathcal{S5}$.

This thesis is structured in the following way. In the first chapter that follows, Chapter 2, we will introduce Gentzen's sequent calculus and the influential Hauptsatz, the cut-elimination theorem. In Chapter 3, we will describe important consequences of the cut-elimination theorem such as Gentzen's Midsequent Theorem, Maehara's Lemma and one of the main theorems in this thesis: Craig's Interpolation Theorem. We will also delve into Beth's Definability Theorem and its relation to Craig's Interpolation Theorem. We will proceed by introducing modal logic and in particular quantified $\mathcal{S}5$ in Chapter 4. There we will also describe the translation of modal formulas to two-sorted classical first-order logic, a fragment of classical first-order logic that abides much by the same rules. In Chapter 5, we will present Fine's result that Craig's Interpolation Theorem does not hold in quantified $\mathcal{S}5$. Ultimately, in Chapter 6 we will show that prenex formulas in $\mathcal{S}5$ do interpolate, as well as formulas containing weak modalities only.

2 First-order logic

As we mentioned in the introduction, Gentzen's sequent calculus has beneficial properties. Gentzen's cut-elimination theorem can be proven for it, allowing to find proofs of valid sequents that are free of any detours. These properties of Gentzen's sequent calculus will be useful for us when establishing the interpolants in the quantified modal logic $\mathcal{S}5$. Thus, this is the calculus that we will present below, following Gaisi Takeuti's book "Proof Theory" ([Tak87]).

In the first section of this chapter we will introduce some basics. We assume that the reader has some knowledge of first-order logic, but we will go through several definitions in order to clarify the notation used in this thesis. Then we will proceed to introduce Gentzen's sequent calculus for classical first-order logic, called "logistisches klassisches Kalkül" LK, in Section 2.2. We end the chapter with Gentzen's Hauptsatz, the cut-elimination theorem. The useful consequences of this theorem such as Craig's Interpolation Theorem will be introduced in the next chapter.

2.1 Basics

A classical first-order **language** from which we will build our formulas consists of individual constants *a, b, c, ...*, function constants *f, g, h, ...*, predicate constants *R, P, Q, ...*, the 0-place predicate symbol $\top$, and variables *x, y, z, ...*. The language further consists of the logical connectives $\neg, \wedge, \vee, \supset, \forall$ and $\exists$, as well as the auxiliary symbols (,), and , (comma). Sticking with the standard approach in elementary logic, the cardinality of the various kinds of symbols is restricted to $\aleph_0$ with order type ω. In the languages we work with the set of variables is infinite and there is at least one predicate symbol apart from $\top$.

Now we define the terms and formulas in our language.

Definition 2.1. Terms are defined inductively as follows:

1. Every individual constant is a term.
2. Every free variable is a term.
3. If f is a function constant with n argument-places and $t_1, \dots, t_n$ are terms, then $f(t_1, \dots, t_n)$ is also a term.
4. These are all terms.

An **atomic formula** is an expression of the form $\top$ or $R(t_1, \dots, t_n)$, where R is a predicate constant with n argument-places and $t_1, \dots, t_n$ are terms.

Formulas are defined inductively as follows:

1. Every atomic formula is a formula.

2. If A and B are formulas, then $\neg A$, $A \wedge B$, $A \vee B$ and $A \supset B$ are formulas.

3. If A is a formula, a a free variable and x a variable not occuring in A, then the expressions $\forall x A'$ and $\exists x A'$ are formulas, where A' is obtained from A by replacing each occurence of a in A by x.

4. These are all formulas.

We insert parentheses when needed to clarify the meaning of a formula and assume the usual precedence rules for our logical connectives. We write $A \equiv B$ as an abbreviation for $(A \supset B) \wedge (B \supset A)$ and $\bot$ to denote $\neg\top$.

2.2 Gentzen's sequent calculus

Now we will describe Gentzen's sequent calculus LK. It's a calculus that only admits the most basic tautologies in the form $A \supset A$ as axioms, but admits a wide variety of deduction rules that can be used to derive valid formulas. We will use Greek capital letters Γ, Δ, Π and Λ to denote finite and possibly empty sequences of formulas separated by commas, like $A, B, \dots, C$. Gentzen introduced his sequent calculus for a language not including the symbol $\top$. We will present a slightly modified version of LK that includes the predicate symbol $\top$ and according sequents, which Takeuti also introduces and refers to as $LK\#$ ([Tak87] p. 31). Introducing the symbol $\top$ does not change the applicability of the theorems that we will introduce below as their proofs are easily extended.

Definition 2.2. A **sequent** is an expression of the form $\Gamma \to \Delta$, where Γ and Δ are formula sequences as described above. Γ is called the **antecedent** and Δ the **succedent** of the sequent. We call the formulas in Γ and Δ **sequent-formulas**.

According to Gentzen ([Gen35], p.181), a sequent $A_1, \dots, A_m \to B_1 \dots, B_n$ has the same meaning as the formula $A_1 \wedge \dots \wedge A_m \supset B_1 \vee \dots \vee B_n$. Consequently, the sequent $A_1, \dots, A_m \to$ expresses that $A_1 \wedge \dots \wedge A_m$ yields a contradiction, and $\to B_1, \dots, B_n$ means that $B_1 \vee \dots \vee B_n$

holds. The empty sequent $\rightarrow$ expresses a contradiction. We will denote sequents by the letter S, with or without subscripts.

Definition 2.3. An **inference** is an expression of the form

$$\frac{S_1}{S} \qquad \text{or} \qquad \frac{S_1 \qquad S_2}{S},$$

where S_1 and S_2 are called **upper sequents** and S the **lower sequent** of the inference.

Intuitively, an inference means that we can infer the lower sequent from the upper one(s).

We will now introduce the inferences that are allowed in Gentzen's sequent calculus. As previously, Greek capital letters denote sequences of formulas as before and A, B, C, D, $F(x)$ denote formulas.

1. **Structural Rules:**

   (a) *Weakening:*

   $$\text{left: } \frac{\Gamma \rightarrow \Delta}{D,\, \Gamma \rightarrow \Delta}, \qquad \text{right: } \frac{\Gamma \rightarrow \Delta}{\Gamma \rightarrow \Delta,\, D},$$

   where D is called the **weakening formula**.

   (b) *Contraction:*

   $$\text{left: } \frac{D,\, D,\, \Gamma \rightarrow \Delta}{D,\, \Gamma \rightarrow \Delta}; \qquad \text{right: } \frac{\Gamma \rightarrow \Delta,\, D,\, D}{\Gamma \rightarrow \Delta,\, D}.$$

   (c) *Exchange:*

   $$\text{left: } \frac{\Gamma,\, C,\, D,\, \Pi \rightarrow \Delta}{\Gamma,\, D,\, C,\, \Pi \rightarrow \Delta}; \qquad \text{right: } \frac{\Gamma \rightarrow \Delta,\, C,\, D,\, \Lambda}{\Gamma \rightarrow \Delta,\, D,\, C,\, \Lambda}.$$

   (d) *Cut:*

   $$\text{left: } \frac{\Gamma \rightarrow \Delta,\, D \qquad D,\, \Pi \rightarrow \Lambda}{\Gamma,\, \Pi \rightarrow \Delta, \Lambda},$$

   where D is called the **cut formula**.

2. **Logical Rules:**

   (a) **Propositional Inferences:**

   i. $\neg$*:*

   $$\text{left: } \frac{\Gamma \rightarrow \Delta,\, D}{\neg D,\, \Gamma \rightarrow \Delta}; \qquad \text{right: } \frac{D,\, \Gamma \rightarrow \Delta}{\Gamma \rightarrow \Delta,\, \neg D}.$$

   ii. $\wedge$*:*

   $$\text{left: } \frac{C,\, \Gamma \rightarrow \Delta}{C \wedge D,\, \Gamma \rightarrow \Delta} \quad \text{and} \quad \frac{D,\, \Gamma \rightarrow \Delta}{C \wedge D,\, \Gamma \rightarrow \Delta};$$

$$\text{right:} \quad \frac{\Gamma \to \Delta,\, C \qquad \Gamma \to \Delta,\, D}{\Gamma \to \Delta,\, C \wedge D}\ .$$

iii. $\vee$:

$$\text{left:} \quad \frac{C,\, \Gamma \to \Delta \qquad D,\, \Gamma \to \Delta}{C \vee D,\, \Gamma \to \Delta}\ ;$$

$$\text{right:} \quad \frac{\Gamma \to \Delta,\, C}{\Gamma \to \Delta,\, C \vee D} \quad \text{and} \quad \frac{\Gamma \to \Delta,\, D}{\Gamma \to \Delta, C \vee D}\ .$$

iv. $\supset$:

$$\text{left:} \quad \frac{\Gamma \to \Delta,\, C \qquad D,\, \Pi \to \Lambda}{C \supset D,\, \Gamma,\, \Pi \to \Delta,\, \Lambda}\ ;$$

$$\text{right:} \quad \frac{C,\, \Gamma \to \Delta,\, D}{\Gamma \to \Delta,\, C \supset D}\ .$$

(b) **Quantifier Inferences:**

i. $\forall$:

$$\text{left:} \quad \frac{F(t),\, \Gamma \to \Delta}{\forall x F(x),\, \Gamma \to \Delta}\ , \qquad \text{right:} \quad \frac{\Gamma \to \Delta, F(y)}{\Gamma \to \Delta, \forall x F(x)}\ ,$$

where t is an arbitrary term, and the variable y does not appear in the lower sequent of $\forall$: *right*. In that case, y is called the **eigenvariable** of the inference and the mentioned condition the **eigenvariable condition**.

ii. $\exists$:

$$\text{left:} \quad \frac{F(y),\, \Gamma \to \Delta}{\exists x F(x),\, \Gamma \to \Delta}\ , \qquad \text{right:} \quad \frac{\Gamma \to \Delta, F(t)}{\Gamma \to \Delta, \exists x F(x)}\ ,$$

where the variable y does not appear in the lower sequent of $\exists$: *left* and t is an arbitrary term. Again, y is called the **eigenvariable** and the mentioned condition the **eigenvariable condition** of this inference.

The first three rules, i.e. the structural rules "Weakening", "Contraction" and "Exchange", are referred to as **weak inferences**, while all others are called **strong inferences**. Further, we refer to the quantifier inferences $\forall : right$ and $\exists : left$, which require the eigenvariable condition to be fulfilled, as *strong* quantifier inferences, and the quantifiers as *strong* quantifiers. The quantifiers introduced with $\forall : left$ and $\exists : right$ are referred to as *weak* in contrast. The condition placed on quantified formulas in Definition 2.1 prohibits the introduction of quantifiers in inferences that are already present in the formula, such that an expression like $\forall x \forall x A(x) \wedge B(x)$ cannot be inferred as it is not a formula.

In addition to the inference rules we define the sequents from which we can start our inferences, so-called **initial sequents** or **axioms**. These are sequents of the form $A \to A$ as well as the sequent $\to \top$.

Definition 2.4. A **proof** is a tree of sequents that satisfies these conditions:

1. The topmost sequents of the tree are initial sequents.

2. Every sequent in the tree except the lowest one is an upper sequent of an inference whose lower sequent is also in the proof.

Definition 2.5. The lowest sequent of a proof P is called **end-sequent**. A sequent S is called **provable** if there is a proof P with S as its end-sequent. A formula A is called **provable** if the sequent $\to A$ is provable.

When we deal with proofs, we will assume that all eigenvariables appearing in the proof are distinct from another, and if a free variable y appears as an eigenvariable in a sequent S of the proof, then it only occurs in sequents in the branch leading to S. According to Lemma 2.10 in [Tak87] we can make this assumption without loss of generality.

Proposition 2.6. *Every sequent that is provable has a proof in which all the initial sequents consist of atomic formulas. If the sequent is provable without the cut-rule (which is always the case in LK as we will see later), then it has a proof without a cut in which all the initial sequents consist of atomic formulas.*

Proof. As all initial sequents of proofs are by definition of the form $\to \top$ or $A \to A$ for some arbitrary formula A, it suffices to show the proposition for the sequents $\to \top$ and $A \to A$. The case for $\to \top$ is clear as $\top$ is an atomic formula. The latter can be shown by induction on the complexity of A. □

Now we have prepared the ground for Gentzen's well-known Hauptsatz and its far-reaching implications.

Theorem 2.7 (Gentzen's Hauptsatz: the cut-elimination theorem)**.** *If a sequent is provable, then it has a proof without a cut, i.e. a cut-free proof.*

Proof. We will present an outline of the proof. Gentzen proved the cut-elimination theorem by defining a new inference rule, the *mix*, and showing that it is equivalent to the cut-rule. Let A be a formula, and let Π' and Δ' denote the sequences Π and Δ with all occurences of A deleted, respectively. Then the mix rule with respect to the formula A is the following inference rule:

$$\frac{\Gamma \to \Delta \qquad \Pi \to \Lambda}{\Gamma, \Pi' \to \Delta', \Lambda} \quad (\text{Mix}, A)$$

After proving the equivalence of the mix and cut rule, Gentzen proceeds to show by transfinite induction that the mix can be eliminated in any proof. The proof is constructive and gives an alternative, cut-free proof of the endsequent. The full proof can be found in [Tak87] on pp. 21-28. □

2.3 Semantics of classical first-order logic

So far we have only talked about the syntactical aspects of classical first-order logic - it is time to explore the semantics. A model in the language of classical first-order logic as we have described it above is a pair $\mathfrak{M} = (D, V)$, where D is a non-empty set, and V is a function mapping all the constants of the language such that $V(a) \in D$ for individual constants a, $V(f) \subset D^n$ for a function constant f with n arguments, and $V(P) \subset D^n$ for a predicate constant P with n arguments. The satisfaction relation that tells us when a model satisfies a formula is defined in the usual manner. Accordingly, a formula is called *valid* if it is satisfied by every model with every variable assignment.

Theorem 2.8 (completeness and soundness)**.** *A formula is provable in LK if and only if it is valid.*

For the proof we refer to [Tak87], p. 40ff.

In place of the algorithmic procedure presented in the proof of the cut-elimination theorem above, we could also prove the Hauptsatz semantically by showing that the cut-free fragment of LK is complete. That is, any valid sequent in LK is also provable in the fragment of LK without a cut. However, this proof is not constructive. Calculus systems for which the non-constructive type of proof exists (but possibly not the constructive proof) are called *cut-free complete* ([Bur], p. 13). As already mentioned in the introduction of this thesis, the cut-elimination theorem is a very important and useful property of Gentzen's sequent calculus. We will explore some theorems that can be proved thanks to cut-elimination in the next chapter, in particular Craig's famous interpolation theorem.

3 Craig's Interpolation Theorem

The cut-elimination theorem has a lot of useful consequences that we will examine in this chapter. One of the theorems that can be proven thanks to Gentzen's Hauptsatz is Craig's Interpolation Theorem. This theorem was originally thought of by William Craig as a lemma in order to prove Beth's Definability Theorem (see Section 3.2) in a simpler way than had been achieved until then. However, it has gained a standing of its own afterwards. We will present some theorems following from the cut-elimination theorem in Section 3.1, orienting ourselves on [Tak87] once again unless indicated otherwise. The exploration culminates in Craig's Interpolation Theorem 3.6. In Section 3.2 we will explore Beth's Definability Theorem and its meaning, following the respective entry in the Routledge Encyclopedia of Philosophy ([Swi98]).

3.1 Consequences of the cut-elimination theorem

The cut-elimination theorem gives us information on the structure of proofs and their end result. We will use the term *subformula* of a formula A to refer to the formulas that are used to build A. If we look at the inferences that are allowed in LK, we can see that every formula appearing in an upper sequent is a subformula of some formula appearing in the lower sequent - except for in the cut-rule. Therefore, in a proof without a cut we only use subformulas of the end-sequent-formulas.

Theorem 3.1 (Subformula property)**.** *All formulas which occur in a cut-free proof in LK are subformulas of some formulas of the proof's end-sequent.*

Proof. The result is shown by induction on the number of inferences in the cut-free proof. □

This is a remarkable property of cut-free proofs. Thus, in such proofs we can infer the end-sequent from the initial sequents without any detours or expressions that are not used in the end-sequent. The only downside of this proof format is that it comes at the cost of readability of the proof, as compared to proofs in Gentzen's natural deduction system for example.

We refer to sequent calculi as *analytic* in which cut-elimination or at least cut-free completeness as well as the subformula property hold. This term was introduced in the 1960s by Smullyan and expresses the idea that the given formula is analyzed in the proof procedure. Proofs in analytic systems have the remarkable property that they can be transformed into proofs without any detours, that is, they can be built from the bottom up consisting only of subformulas of the end-sequent. Systems like Hilbert-style systems or Gentzen's natural deduction system are *non-analytic.* In those systems, usually either the cut-rule or Modus Ponens are used to eliminate formulas which are not part of the end-sequent. It is possible to have an analytic calculus that uses the cut-rule, however, as the subformula property may hold nevertheless ([Pfe84]).

A direct consequence of the subformula property is the consistency of LK.

Theorem 3.2. *Gentzen's sequent calculus for classical first-order logic, LK, is consistent.*

Proof. Suppose LK were inconsistent, i.e. the empty sequent $\rightarrow$ could be derived in it. By the cut-elimination theorem, this would be provable without using the cut-rule. However, this is impossible by the subformula property. □

Gentzen also introduced an important theorem about the structure of proofs of prenex formulas, the Midsequent Theorem that we will describe below.

Definition 3.3. A formula is called **prenex** if no quantifier in it is in the scope of a propositional connective. A sequent is called **prenex sequent** if it consists of prenex formulas only.

Every formula in LK is equivalent to a prenex formula, that is, for every formula A there exists a prenex formula B such that $A \equiv B$ is deducible in LK. According to the Midsequent Theorem, we can find a proof for a sequent consisting of prenex formulas that is divided into two parts - the propositional part, where only inference rules that are also applicable in propositional logic are applied, and the quantificational part, in which quantifiers are introduced. Since we will focus on prenex formulas and their interpolant in the context of the modal logic $\mathcal{S}5$ later in the thesis, the Midsequent Theorem will play a central role in later chapters.

Theorem 3.4 (Gentzen's Midsequent Theorem)**.** *For a provable prenex sequent S, there is a cut-free proof which has the following property: The proof contains a quantifier-free sequent M called the* ***mid-sequent*** *such that every inference above M is either structural or propositional, and every inference below it is either structural or a quantifier inference.*

Proof. We will give an outline of the proof. Let S be a provable sequent which only consists of formulas in prenex form. From Proposition 2.6 and the cut-elimination theorem we know that there is a cut-free proof P of S in which all initial sequents consist of atomic formulas only. We will define an *order of P* and conduct the proof as an induction on this order. Let I be a quantifier inference in P. Then we define the *order of I* as the number of propositional inferences that appear in P below I. The *order of P* thus is defined as the sum of the orders of all quantifier inferences in P.

We start with the base case where the order of P is 0. In this case, if there is a propositional inference in P, then there is no quantifier inference above it. We take the lowest propositional inference and refer to its lower sequent as M_0. Even though there is no quantifier inference above it, M_0 might still contain quantifiers introduced by weakenings. By the subformula property, these weakening formulas must appear as subformulas in the formulas of the endsequent. These formulas are in prenex form, such that in the course of the proof, no propositional inferences are applied to the weakening formulas containing quantifiers. Thus, we can just eliminate these weakenings above M_0 and if necessary introduce them again below it. This way we get an adapted sequent M_0' that serves as a quantifier-free midsequent for the proof of S.

If there is no propositional inference in P, we can take the upper sequent of the uppermost quantifier inference as the midsequent.

In the case that the order of P is larger than 0, we can find at least one propositional inference under a quantifier inference. In particular, we can find a quantifier inference I with the property that the uppermost logical (i.e. quantifier or propositional) inference is a propositional inference I'. Now we can lower the order of P by interchanging the positions of the inferences I and I'. For example, in the case of I being $\forall : \mathit{right}$, the proof P has this form:

$$\begin{array}{lc}
 & \cdots \\ \hline
I & \Gamma \to \Pi, F(y) \\ \hline
 & \Gamma \to \Pi, \forall x F(x) \\ \hline
I' & \text{structural inferences} \\ \hline
 & \Delta \to \Lambda
\end{array}$$

where Λ contains $\forall x F(x)$ as a sequent-formula. We can transform P into the following proof P' by interchanging the positions of I and I':

$$\begin{array}{lc}
 & \Gamma \to \Pi, F(y) \\ \hline
 & \text{structural inferences} \\ \hline
 & \Gamma \to F(y), \Pi, \forall x F(x) \\ \hline
I' & \cdots \\ \hline
 & \Delta \to F(y), \Lambda \\ \hline
I & \cdots \\ \hline
 & \Delta \to \Lambda, \forall x F(x) \\ \hline
 & \cdots \\ \hline
 & \Delta \to \Lambda \\ \hline
 & \cdots
\end{array}$$

Clearly, the proof P' has a lower order than P. □

Now we introduce Maehara's partition interpolation method. It is the basis of the proof of Craig's Interpolation Theorem.

Lemma 3.5 (Maehara's Lemma)**.** *Let $\Gamma \to \Delta$ be an LK-provable sequent. Further, let (Γ_1, Γ_2) and (Δ_1, Δ_2) be arbitrary partitions of Γ and Δ respectively. We will denote this partition of the sequent $\Gamma \to \Delta$ as $[\{\Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$. Then there exists a formula C, called the interpolant of the partition, such that:*

1. *$\Gamma_1 \to \Delta_1, C$ and $C, \Gamma_2 \to \Delta_2$ are both provable in LK;*
2. *C only contains free variables and individual and predicate constants (apart from $\top$) that occur in $\Gamma_1 \cup \Delta_1$ as well as $\Gamma_2 \cup \Delta_2$.*

Before we prove Maehara's Lemma, we will derive Craig's Interpolation Theorem from it.

Theorem 3.6 (Craig's Interpolation Theorem)**.** *Let A and B be two formulas such that $A \supset B$ is provable in LK. Then there is a formula C, the* ***interpolant*** *of A and B, such that $A \supset C$ and $C \supset B$ are provable in LK, and such that C only contains free variables and individual and predicate constants (apart from $\top$) that occur in A as well as B.*

Proof. (Craig's Interpolation Theorem). Let A and B be two formulas such that $A \supset B$ is provable in LK. The sequent $A \to B$ is provable then, too, and we can apply Maehara's Lemma using the partition $[\{A;\}, \{;B\}]$. Consequently, there exists a formula C satisfying the conditions 1. and 2. of the lemma. That is, the sequents $A \to C$ and $C \to B$ are provable in LK, and C only contains free variables and individual and predicate constants (apart from $\top$) that occur in A as well as B. Thus, C is the required interpolant of $A \supset B$.

□

The proof of Craig's Interpolation Lemma is based on Maehara's Lemma. In the proof thereof, we derive an interpolant for the given sequent constructively by building on the logical structure of the proof. The interpolant only contains the information truly necessary to maintain the implication. Thus, the interpolant of an implication encapsulates its logical essence.

Proof. (Maehara's Lemma). The lemma is proven by induction on the number of inferences k in a cut-free proof of the sequent $\Gamma \to \Delta$. We will focus on the cases of inferences that we will also use later for our method to interpolate formulas in the modal logic $\mathcal{S}5$. We will also restrict the partitions we look at to the one relevant for our purpose, where $\Gamma_1 = \Gamma$, $\Delta_2 = \Delta$, and Γ_2

and Δ_1 are both empty, i.e. the partition $[\{\Gamma;\}, \{;\Delta\}]$ used in the proof of Craig's Interpolation Theorem. Thus we look for interpolants C such that $\Gamma \to C$ and $C \to \Delta$ are provable.

1. $k = 0$: First we start with the base case where the number of inferences is $k = 0$. Thus, our sequent $\Gamma \to \Delta$ has the form of an initial sequent $\to \top$ or $D \to D$. In the first case, the interpolant is $\top$. In the second case, we look at the partition $[\{D;\}, \{;D\}]$. The formula D fulfills all the requirements for an interpolant of $D \to D$.

2. $k > 0$ and the last inference is $\forall: left$:

$$\frac{F(t), \Gamma \to \Delta}{\forall x F(x), \Gamma \to \Delta}$$

Let $b_1, \ldots, b_n$ denote the free variables and constants (possibly none) in the term t. Again, we suppose our partition is $[\{\forall x F(x), \Gamma;\}, \{;\Delta\}]$. This induces the partition $[\{F(t), \Gamma;\}, \{;\Delta\}]$ in the upper sequent. We can apply the induction hypothesis to find an interpolant $C(b_1, \ldots, b_n)$ such that

$$F(t), \Gamma \to C(b_1, \ldots, b_n) \qquad \text{and} \qquad C(b_1, \ldots, b_n) \to \Delta$$

are both provable in LK. Now let $b_{i_1}, \ldots, b_{i_m}$ be all the variables and constants among $b_1, \ldots, b_n$ that do not occur in $\{F(x), \Gamma\}$. Since the interpolant may only contain free variables and constant symbols that occur in both the antecendent and succedent of the sequent, we need to replace $b_{i_1}, \ldots, b_{i_m}$ with bound variables. Then the required interpolant $\tilde{C}$ has the form

$$\forall y_1 \ldots \forall y_m C(b_1, \ldots, y_1, \ldots, y_m, \ldots, b_n),$$

where we replaced $b_{i_1}, \ldots, b_{i_m}$ by the bound variables $y_1, \ldots, y_m$.

We can derive that $\tilde{C}$ fulfills the requirements that $\forall x F(x), \Gamma \to \tilde{C}$ and $\tilde{C} \to \Delta$ are LK-provable with the following inferences:

$$\forall\text{: right}\ \frac{\forall\text{: left}\ \dfrac{F(t), \Gamma \to C(b_1, \ldots, b_n)}{\forall x F(x), \Gamma \to C(b_1, \ldots, b_n)}}{\forall x F(x), \Gamma \to \forall y_1 \ldots \forall y_m C(b_1, \ldots, y_1, \ldots, y_m, \ldots, b_n)}$$

$$\forall\text{: left}\ \frac{C(b_1, \ldots, b_n) \to \Delta}{\forall y_1 \ldots \forall y_m C(b_1, \ldots, y_1, \ldots, y_m, \ldots, b_n) \to \Delta}$$

3. $k > 0$ and the last inference is $\forall: right$:

$$\frac{\Gamma \rightarrow \Delta, F(y)}{\Gamma \rightarrow \Delta, \forall x F(x)} \ ,$$

where y is a free variable which does not occur in the lower sequent. We partition the lower sequent as $[\{\Gamma;\}, \{;\Delta, \forall x F(x)\}]$, which induces the partition $[\{\Gamma;\}, \{;\Delta, F(y)\}]$ on the upper sequent. By our induction hypothesis, there is an interpolant C of the upper sequent which fulfills all the requirements of Maehara's Lemma, in particular the provability of the implications $\Gamma \rightarrow C$ and $C \rightarrow \Delta, F(y)$. Since the variable y does not occur in $\{\Gamma, \Delta, F(x)\}$, it does not appear in C either and we can infer:

$$\forall\text{: right} \quad \frac{C \rightarrow \Delta, F(y)}{C \rightarrow \Delta, \forall x F(x)}$$

Thus, C is also an appropriate interpolant for the lower sequent $\Gamma \rightarrow \Delta, \forall x F(x)$.

4. $k > 0$ and the last inference is $\exists: left$:

$$\frac{F(y), \Gamma \rightarrow \Delta}{\exists x F(x), \Gamma \rightarrow \Delta} \ ,$$

This case is analogous to the one of the inference $\forall: right$ that we just described above. Therefore, the interpolant C of the upper sequent is also suitable interpolant of the lower sequent of the inference.

5. $k > 0$ and the last inference is $\exists: right$:

$$\frac{\Gamma \rightarrow \Delta, F(t)}{\Gamma \rightarrow \Delta, \exists x F(x)}$$

This case is analogous to the one of the inference $\forall: left$. Let $b_1, \ldots, b_n$ denote all the free variables and constants (possibly none) in the term t and let $C(b_1, \ldots, b_n)$ be the interpolant of the upper sequent. Then $\exists y_1 \ldots \exists y_m C(b_1, \ldots, y_1, \ldots, y_m, \ldots, b_n)$ is a suitable interpolant for the lower sequent, where the y_i replaced all the elements among $b_1, \ldots, b_n$ that do not occur in $\{\Delta, F(x)\}$.

Let us describe the interpolant for other inferences as well for the sake of completeness.

The structural inferences *Weakening*, *Contraction* and *Exchange* do not affect the logical consequence of a sequent. Thus, the interpolant of the lower sequent is the same as the one from the upper sequent.

For the propositional inferences, the interpolant remains either unchanged or is a propositional connection of the interpolants of the upper sequents. For example, for the inference rules $\wedge : left$, the interpolant of the antecedent is the same as for the succedent, as introducing the additional formula this way has no effect on the essence of the implication, just was with a weakening. On the other hand, let us consider the inference rule $\wedge :\ right$:

$$\frac{\Gamma \to \Delta,\, C \qquad \Gamma \to \Delta,\, D}{\Gamma \to \Delta,\, C \wedge D}\ .$$

If the left and right upper sequents have the interpolant C_1 and C_2 respectively, then the lower sequent has the interpolant $C_1 \wedge C_2$. □

As we have seen in the proof of Maehara's Lemma, the interpolant of an implication $A \supset B$ can be formed constructively by looking at the inferences of the proof and forming the interpolant in an inductive manner from the bottom up. If we deal with prenex formulas in particular, we can apply the Midsequent Theorem (Theorem 3.4) first and obtain a proof where all quantifier inferences only appear below the midsequent. The midsequent is then quantifier-free and we can obtain its interpolant using the methods of classical propositional logic. We can then use the inductive steps of the proof of Maehara's Lemma to reintroduce the quantifiers to the midsequent's interpolant to obtain the interpolant of the original formula. Thus, the combination of the Midesequent Theorem and Maehara's Lemma provides us with tools to constructively build the interpolant of prenex formulas in a straight-forward manner. We will further explore this approach later in the thesis.

3.2 Beth's Definability Theorem

Beth's Definability Theorem is a central result in classical first-order logic about the definability of non-logical symbols. We will explore the context of this important theorem in this section following the *Routledge Encyclopedia of Philosophy* [Swi98]. The question in which ways concepts can be defined began to arise with the development of the formal axiomatic method in the mathematics of the late nineteenth and twentieth century. Beth's theorem as described in [Fin79] states that in classical first-order logic, a predicate symbol P is implicitly definable in a theory T if and only if it is explicitly definable in T:

Definition 3.7. Let T be a theory and T' be the result of replacing each occurence of the predicate symbol P in T by a new predicate symbol P' of the same degree $n \geq 0$. Then we say that P is **implicitly definable** in T if $T, T' \vdash \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \supset P'(x_1, \dots, x_n))$.

We say that P is **explicitly definable** in T if there is a formula φ in the language of T not containing P such that $T \vdash \forall x_1, \ldots \forall x_n (P(x_1, \ldots, x_n) \equiv \varphi)$.

Theorem 3.8 (Beth's Definability Theorem)**.** *A predicate P is implicitly definable in a theory T if and only if it is explicitly definable in T.*

Implicit definability can intuitively be described as the theory T fixing the extension of the symbol P uniquely if the extension of the other symbols in the language is given. This translates to conditions being placed on the models of T, that is, to semantical conditions. On the contrary, explicit definability of P means that its extension can be explicitly given by a formula, i.e. it expresses a syntactical condition as it needs to be deducible that P is equivalent to this formula. Consequently, Beth's theorem tells us something about the expressive power of classical first-order logic - it tells us that there is a balance between the semantics and the syntax of the logic at hand.

The expressions of "implicit" and "explicit" definability of concepts originates in the early nineteenth century. The French mathematician Jose Diez Gergonne suggested these terms inspired by their use in algebra. There they were used to describe unknowns that are defined by an unsolved or solved set of equations respectively. In 1901, the Italian mathematician Alessandro Padoa proposed a new method, later called "Padoa's Method", to prove the independence of a concept of other concepts in a theory: he aimed at doing so by finding two models that differ only on the term in question. That is, his method entailed establishing that a concept was *not* implicitly definable in a theory. Padoa claimed that this was a sufficient and necessary condition to prove explicit undefinability as well.

Sufficiency is clear, as it means that implicit definability follows from explicit definability. If a predicate P is explicitly defined in a theory T, then for any model satisfying the equivalence of the explicit definition, the meaning of P is already uniquely interpreted due to the explicit definition. Consequently, two models agreeing on every term other than P must also agree on the extension of P.

Necessity is less obvious, however. The question could not have been answered at that time, as it requires the underlying logic to be specified more precisely than what Padoa had at his hands. In 1953, Evert Beth, a Dutch philosopher and logician, managed to prove that implicit definability also implied explicit definability for classical first-order logic – and Beth's Definability Theorem was born. More precisely, Beth showed that if a term cannot be explicitly defined in a theory, then there exist two models of the theory that agree on the interpretation of all terms except for the one in question. Thus the term cannot be implicitly defined either. In his original proof, Beth used Gentzen's Hauptsatz, the cut-elimination theorem (Theorem 2.7) that we have introduced above, and he was even able to find a way to construct two models that differ on the term in question explicitly using his so-called semantic tableau method. Today, however, Beth's Definability Theorem is usually shown as a direct implication of Craig's Interpolation Theorem (Theorem 3.6). This theorem was in fact originally introduced by William Craig as a lemma in order to prove Beth's Definability Theorem in a simpler way, but has gained such attention ever since that today it stands on its own.

We already presented in the preceding paragraphs why implicit definability follows from explicit definability. We will now sketch the rest of the proof of Beth's Definability Theorem, that is, the direction that implicit definability implies explicit definability, using Craig's Interpolation Theorem. Let us assume that the predicate symbol P is implicitly definable in the theory T. That is, if T' is the theory obtained from T by replacing every occurrence of P by a new predicate symbol P' of the same degree, we have that

$$T, T' \vdash \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \supset P'(x_1, \dots, x_n)).$$

By compactness, we can assume T and T' to be a finite set of sentences that we can write as a conjunction. Thus we can say that

$$T \wedge T' \vdash \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \supset P'(x_1, \dots, x_n)).$$

Consequently, the following deduction also holds:

$$T \wedge P(c_1, \dots, c_n) \vdash T' \supset P'(c_1, \dots, c_n)).$$

where the c_i are new individual constants not contained in the first-order language used. Now we apply Craig's Interpolation Theorem. Accordingly, there is an interpolant $\phi(c_1, \dots, c_n)$ such that $T \wedge P(c_1, \dots, c_n) \vdash \phi(c_1, \dots, c_n)$ and $\phi(c_1, \dots, c_n) \vdash T' \supset P'(c_1, \dots, c_n)$. Since ϕ only contains symbols of the common language of the interpolated formulas (apart from possibly $\top$), the symbol P' is not contained in ϕ. Thus it holds equally that $\phi(c_1, \dots, c_n) \vdash T \supset P(c_1, \dots, c_n)$. From these statements we can deduce the equivalence

$$T \vdash P(c_1, \dots, c_n) \equiv \phi(c_1, \dots, c_n)$$

Thus, P is also defined explicitly in T and we have completed the proof.

4 Modal logic and quantified S5

In this chapter, we will describe modal logic and the quantified version of the modal system $S5$. Already Aristoteles talked about modal logic, and its semantics was being discussed during medieval times as well. The examination of modal logic was rekindled before the start of World War I by the logician C. I. Lewis, though it was first perceived very critically. This changed when the relational semantics where introduced in the 1960s with decisive contributions by Saul A. Kripke ([Pri08] p.70), as they provided an intuitive way to talk about modal logics.

We will start by introducing basic propositional modal logic and its relational semantics. Then we will proceed to talk about quantified modal logic. We will define propositional and quantified modal logic as described in Chapters 1 & 9 of the *Handbook of Modal Logic* edited by Blackburn et al. ([BvB07], [BG07]). Here, modal logic is introduced from a semantic perspective, i.e. as a tool for talking about structures or models. We will introduce quantified modal logic from the sources mentioned above with the exception that we will not introduce the equality symbol, as we have introduced first-order logic without equality as well. Modal logic is closely linked to classical first-order logic by the standard translation, which enables us to view modal logic as a fragment of the former. We will describe two-sorted logic, which is a fragment of classical first-order logic, and present the standard translation mapping modal formulas to formulas in two-sorted logic.

4.1 Propositional modal logic

4.1.1 Basics

In this section, we will introduce propositional modal logic as described in [BvB07]. The signature of our modal language will be a set of propositional symbols, typically denoted as p, q, $r, \dots$ and the modality symbol $\Box$. We will usually assume a fixed signature where the set of propositional symbols is denumerably infinite. We define the **basic modal language** given a signature as follows:

Definition 4.1. We define **formulas** in propositional modal logic in the following way:

1. Any propositional symbol is a formula.

2. The logical symbols $\top$ and $\bot$ are formulas.

3. If A and B are formulas, then $\neg A$, $A \wedge B$, $A \vee B$ and $A \supset B$ are formulas.

4. If A is a formula, $\Box A$ is a formula. We write $\Diamond A$ as an abbreviation for $\neg\Box\neg A$.

4.1.2 Relational semantics

Now we will look at a way to interpret propositional modal formulas.

Definition 4.2. We define a **model** $\mathfrak{M}$, also called **Kripke model**, for the basic modal language over some fixed signature as a triple $\mathfrak{M} = (W, R, V)$, where

1. W is a non-empty set that we refer to as **domain** and whose elements we will call **worlds**,
2. R is a binary relation on the worlds W, the so-called **accessibility relation**,
3. V is a function called **valuation** assigning to each propositional symbol p a subset $V(p) \subset W$

We can think of the set $V(p)$ as the set of worlds in which p is true. An intuitive way to think of Kripke models is to imagine them as graphs, where the vertices represent the worlds and the (directed) edges stand for the binary relation on the worlds. The drawing in Figure 1 is an example of such a graph. It represents a model $\mathfrak{M} = (W, R, V)$ where $W = \{u, v, w\}$, $R = \{(u, v), (v, v), (v, w), (w, w), (w, v)\}$, $V(p) = \{u, v\}$ and $V(q) = \{u, w\}$.

We give the **satisfaction** definition for formulas in propositional modal logic now and inductively define in which cases a formula φ is satisfied in a model $\mathfrak{M}$ at a world w. We omit some

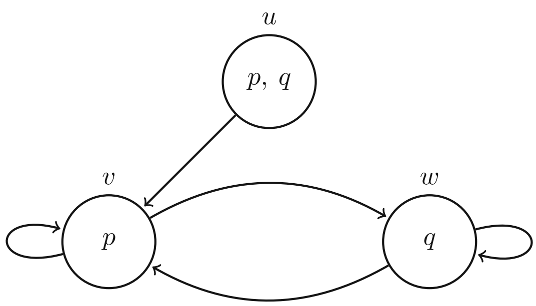


Figure 1: An example of a Kripke model.

Boolean clauses that are clear from analogy.

$$
\begin{array}{rll}
(\mathfrak{M}, w) \models p & \text{iff} & w \in V(p) \\
(\mathfrak{M}, w) \models \top & & \text{always} \\
(\mathfrak{M}, w) \models \bot & & \text{never} \\
(\mathfrak{M}, w) \models \neg\varphi & \text{iff} & (\mathfrak{M}, w) \not\models \varphi \\
(\mathfrak{M}, w) \models \varphi \wedge \psi & \text{iff} & (\mathfrak{M}, w) \models \varphi \text{ and } (\mathfrak{M}, w) \models \psi \\
(\mathfrak{M}, w) \models \Box\varphi & \text{iff} & \text{for all } v \in W \text{ with } R(w, v) \text{ we have } \mathfrak{M}, v \models \varphi \\
(\mathfrak{M}, w) \models \Diamond\varphi & \text{iff} & \text{there is some } v \in W \text{such that } R(w, v) \text{ holds and } \mathfrak{M}, v \models \varphi
\end{array}
$$

If a formula φ is satisfied at all worlds in a model $\mathfrak{M}$, then we call φ **globally satisfied** and write $\mathfrak{M} \models \varphi$. We call φ **valid** if it is globally satisfied in all models and write $\models \varphi$ in that case.

As described in [BvB07] (p.5), the satisfaction definition given above has a very internal character: a formula tells us something about a Kripke model from the inside. A modal formula is always evaluated at a certain world in the model, and thus takes contextual information into account like the propositional variables assigned to this world and connected worlds by a valuation - or in modal first-order logic, which we will introduce below, the variable assignment in that respective world. In first-order classical logic, in contrast, we talk about models from the outside. Here, a closed formula is simply true or false of a given model irrespective of any contextual information regarding the variables.

Furthermore, we can already observe from the satisfaction definition that the modal operators, the box and the diamond, can be seen as an encoding of quantification over the worlds that are accessible via R in a variable-free notation. This will become more apparent with the standard translation that we will present later, which transforms modal formulas to formulas in two-sorted classical first-order logic.

4.1.3 Proof system

Now that we have introduced semantics for basic propositional modal logic, we will present a proof system that we can use to deduce valid formulas. We can axiomatize the set of all modal validities, i.e. the minimal modal logic, by a Hilbert-style proof system $\mathcal{K}$. We define this proof system by declaring the **axioms** of $\mathcal{K}$ to be the following:

1. all formulas in the above basic modal language which have the form of propositional tautologies

2. all instances of the following axiom schema:

$$(K) \qquad \Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi)$$

Furthermore, we may use the following two **inference rules**, *modus ponens* and *modal generalization*:

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

$$(MG) \qquad \frac{\varphi}{\Box\varphi}$$

Remember that validity in the above defined semantics means that a formula is true at *every* world of every model, such that the modal generalization is sound.

4.1.4 The calculus S5

We can examine calculi that contain $\mathcal{K}$ by adding more axioms to our list. One of the most closely studied calculi is the system $\mathcal{S}5$. It is obtained by adding the following axiom schemata to $\mathcal{K}$:

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

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

$$(B) \qquad A \supset \Box\Diamond A$$

The Kripke models satisfying these schemata are exactly those whose relation R is reflexive (T), transitive (4) and symmetric (B). Thus, models in $\mathcal{S}5$ are exactly those which have an equivalence relation as world relation R. We assume without loss of generality that there only is one equivalence class of worlds. Consequently, we have that $R = W \times W$. In light of this relationship, we refer to these kinds of models as $\mathcal{S}5$-models.

Theorem 4.3. *The logic $\mathcal{S}5$ is sound and complete with respect to the Kripke semantics given above for $\mathcal{S}5$-models.*

From the axiom schemata for $\mathcal{S}5$ given above we can infer other formulas valid in $\mathcal{S}5$ (see [Ste88], p. 113), for example:

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

These formulas can also be deduced semantically, as they follow from the simple fact that in $\mathcal{S}5$-models all worlds are connected to another. From this we can see that if we have several modal operators at the beginning of a propositional $\mathcal{S}5$-formula, only the innermost operator is relevant for the formula's meaning.

4.2 Quantified modal logic

Now we move on to the quantified version of modal logic. We will follow [BG07] in this section, though in order to stay consistent with our notation throughout this thesis, we will define models the way that Fine does in his paper [Fin79] that we will refer to in Chapter 5.

4.2.1 Basics

The syntax for the basic first-order modal logic is, analogously to propositional modal logics, obtained simply by taking the syntax of classical first-order logic as we have defined it in Chapter 2 and adding a modal operator $\Box$. Just as in classical logic, a set of predicate symbols is given which we will denote as previously using the metavariables $P, Q, R, \ldots$. For simplicity, we will not consider functions or individual constants. From a model-theoretic perspective, constants are just variables that are not being quantified over. Function symbols can be modelled using predicate symbols. Thus, all **terms** in our modal language will be variables. We will refer to the language of quantified modal logic described here as $\mathcal{QML}$. Now we will proceed to define formulas, models and the satisfaction relation for first-order modal formulas.

Definition 4.4. We define **formulas** in quantified modal logic in the following way:

1. For any predicate symbol P of arity n, the expression $P(x_1, \ldots, x_n)$ is a formula.

2. The logical symbols $\top$ and $\bot$ are formulas.

3. If A and B are formulas, then $\neg A$, $A \wedge B$, $A \vee B$ and $A \supset B$ are formulas.

4. If A is a formula, then $\forall x A$ and $\Box A$ are formulas. We write $\Diamond A$ as an abbreviation for $\neg\Box\neg A$.

Regarding the semantics of quantified modal logic that we chose to present here, the variables designate *rigidly*. This means that the variables designate the same object in every world and we do not need to set a specific domain for every world when defining a model. Predicates, however, may have different extensions in different worlds. We also need to make a choice what domains our quantifiers quantify over in the different worlds. In this thesis, we will work with *constant domain* quantified modal logic. That is, we only need to set one quantification domain D which holds for all worlds when defining the model.

Definition 4.5. A **constant domain model** $\mathfrak{M}$ is a tupel (W, R, D, V), where

1. W is a non-empty set,
2. R is a binary relation on W,
3. D is a non-empty set,
4. V is a function mapping each n-place predicate symbol P to a subset $V(P) \subset W \times D^n$.

As before, we refer to the elements of W as *worlds*, R stands for the *accessibility relation*, the set D is the *domain of quantification*, and V refers to the *valuation*.

Definition 4.6. A **variable assignment** is a function usually referred to as $\mathfrak{a}$ or $\mathfrak{b}$ that assigns an element of the domain D to each variable for a constant domain model $\mathfrak{M} = (W, R, D, V)$.

The **satisfaction relation** is defined as usual with the following definitions for the predicates and quantifiers, where $\mathfrak{a}$ is a variable assignment:

$$\begin{aligned}
(\mathfrak{M}, w) \models P(x_1, \ldots, x_n)[\mathfrak{a}] \quad & \text{iff} \quad (w, \mathfrak{a}(x_1), \ldots, \mathfrak{a}(x_n)) \in V(P) \\
(\mathfrak{M}, w) \models \forall x \varphi[\mathfrak{a}] \quad & \text{iff} \quad (\mathfrak{M}, w) \models \varphi[\mathfrak{b}] \text{ for every assignment } \mathfrak{b} \text{ that differs from } \mathfrak{a} \\
& \qquad \text{at most on } x
\end{aligned}$$

The semantic condition of having constant domains can be expressed syntactically with the Barcan formula $\forall x \Box \phi \supset \Box \forall x \phi$ and its converse $\Box \forall x \phi \supset \forall x \Box \phi$. That is, models satisfying the equivalence $\Box \forall x \phi \equiv \forall x \Box \phi$ are constant domain models.

4.2.2 Hilbert-style proof system for the constant domains

We have already given Hilbert-style axiom systems and inference rules for the propositional modal logics $\mathcal{K}$ and $\mathcal{S5}$ in Section 4.1.3. We will now expand them to cover the quantified logics as well. To receive the axiom system for constant domain $\mathcal{K}$, we expand the propositional axiom scheme by the Barcan formula mentioned in the previous section and the ($\forall$*Elimination*)-axiom:

$$\forall x \Box \varphi(x) \equiv \Box \forall x \varphi(x) \qquad \textit{(Barcan)}$$

$$\forall x \varphi(x) \supset \varphi[y/x] \qquad \textit{(}\forall\textit{ Elimination)}$$

as well as the inference rule

$$\frac{\psi \supset \varphi[y/x]}{\psi \supset \forall x \varphi(x)} \qquad (\forall\ \textit{Introduction})$$

where the rule ($\forall$*Introduction*) has the variable condition that y does not occur freely in the lower implication. The expression $\varphi[y/x]$ refers to the formula φ with all occurrences of x replaced by y. The axiom system for quantified constant-domain $\mathcal{S5}$ is given by taking the system for propositional $\mathcal{S5}$ described in Section 4.1.4 and adding the two axioms and the rule just given. This system is sound and complete with respect to the constant domain semantics given in Section 4.2.1 for constant domain models having the accessibility relation $R = W \times W$ ([BG07] p. 555). We will refer to constant-domain $\mathcal{S5}$ as $\mathcal{S5B}$ from now on.

4.2.3 Sequent calculus for quantified modal logics

In their paper "Gentzen Method in Modal Calculi" from 1957, Masao Ohnishi and Kazuo Matsumoto describe a Gentzen-style sequent calculus for propositional modal logics ([OM57]). This sequent calculus is analogous to Gentzen's sequent calculus LK, omitting the quantifier inferences and adding modal inferences to accomodate the modal operator $\Box$. We can describe such a sequent calculus for quantified modal logic by expanding Gentzen's sequent calculus LK by the modal inferences and leaving the quantifier inferences. A sequent S is defined for modal formulas analogously to sequents in LK. We use the same notation as for LK, with the addition that we write $\Box\Gamma$ to denote the expression $\Box A_1, \dots, \Box A_n$ for Γ being a sequence of formulas $A_1, \dots, A_n$. The following inference rules can be added to the sequent calculus to accomodate different modal logics ([Bur], p.20):

$$\frac{\Gamma \to A}{\Box\Gamma \to \Box A} \qquad \text{(k)}$$

$$\frac{A,\ \Gamma \to \Delta}{\Box A,\ \Gamma \to \Delta} \quad (t)$$

$$\frac{\Gamma \to \Box\Delta,\ A}{\Box\Gamma \to \Delta, \Box A} \quad (b)$$

$$\frac{\Gamma \to}{\Box\Gamma \to} \quad (d)$$

$$\frac{\Box\Gamma \to A}{\Box\Gamma \to \Box A} \quad (4)$$

$$\frac{\Box\Gamma \to \Box\Delta,\ A}{\Box\Gamma \to \Box\Delta,\ \Box A} \quad (5)$$

For the most basic system $\mathcal{K}$, we add the rule (k) only. For the system $\mathcal{S}5$, we further add the rules (t) and (5) or alternatively (d), (b) and (4). Just as in the sequent calculus for classical first-order logic, the initial sequents will be sequents of the form $A \to A$ and $\to \top$.

Ohnishi and Matsumoto prove that the propositional sequent calculus is cut-free for $\mathcal{K}$. However, for $\mathcal{S}5$ it is not cut-free complete. For example, the propositional sequent $A \to \Box\Diamond A$ is not derivable without the use of the cut-rule ([Bur], p.20):

$$\dfrac{\neg:right\ \dfrac{(t)\ \dfrac{\neg:left\ \dfrac{A \to A}{\neg A,\ A \to}}{\Box\neg A, A \to}}{A \to \neg\Box\neg A} \qquad \neg:left\ \dfrac{(5)\ \dfrac{\neg:right\ \dfrac{\Box\neg A \to \Box\neg A}{\to \Box\neg A,\ \neg\Box\neg A}}{\to \Box\neg A,\ \Box\neg\Box\neg A}}{\neg\Box\neg A \to \Box\neg\Box\neg A}}{A \to \Box\neg\Box\neg A}\ Cut$$

Without the cut rule, we could only derive this sequent using weakening and contraction as structural rules. Thus, a derivation of this sequent could only contain sequents of the form:

$$\underbrace{A, \ldots, A}_{m-times} \to \underbrace{\Box\neg\Box\neg A, \ldots, \Box\neg\Box\neg A}_{n-times}$$

with $m, n \geq 0$. Hence, the derivation could not contain an initial sequent. ([LR15]).

Thus, the sequent calculus is not analytic for propositional $\mathcal{S}5$. Consequently, the quantified version is not analytic either for quantified $\mathcal{S}5$.

4.3 Translation to two-sorted first-order logic

As is apparent from the definition of modal logic and its semantics, the modal symbol $\Box$ very much behaves like the quantifier $\forall$ in classical first-order logic. This parallel becomes more obvious with the standard translation that transforms modal formulas to classical ones. The standard translation shows that propositional and first-order modal logic can be regarded as a fragment of classical first-order logic. When reduced to propositional modal logic, we can see that this logic is simply a variable-free notation for a fragment of first-order logic ([BvB07], p. 10). We will first explore two-sorted logic itself before introducing the standard translation.

4.3.1 Two-sorted logic

The standard translation translates any formula in basic first-order modal logics to formulas in two-sorted logic. Before giving the standard translation, we define the two-sorted logic under consideration. The logic being *two-sorted* means that there are two kinds of variables and individual constants – those of the sort world and those of the sort domain for individuals. Two-sorted logic is then a fragment of classical first-order logic. Formulas in two-sorted logic can easily be embedded in classical first-order logic - we just need to add extra predicates to indicate whether a variable should be of the sort world or the sort domain. We will follow the definition of two-sorted logic and standard translation given in [SW00].

We start by defining the language $\mathcal{SL}$ for two-sorted logic. We have variables x, y, $z, \dots$ as variables of the domain sort, and variables u, v, $w, \dots$ to denote the variables of world sort. These two sets of variables are both countably infinite and disjoint. We have predicate symbols $P', Q', R' \dots$ analogously to the predicate symbols in modal first-order logic, such that if P is an n-place predicate in modal logic, then P' denotes an $n+1$-place predicate in two-sorted logic of the form $\mathit{world} \times \mathit{domain}^n$. Again, there are no function symbols or constants. Formulas in $\mathcal{SL}$ are defined in the following manner:

Definition 4.7. We define **formulas** in two-sorted modal logic as:

1. Atomic formulas $P'(v, x_1, \dots, x_n)$ for a predicate symbol P' of sort $\mathit{world} \times \mathit{domain}^n$, and the atomic formula $\top$ are formulas.

2. If φ and ψ are formulas, then $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$ and $\varphi \supset \psi$ are formulas.

3. If φ is a formula, a a free variable of domain sort and x a domain variable not occuring in φ, then the expressions $\forall x \varphi'$ and $\exists x \varphi'$ are formulas, where φ' is obtained from φ by replacing each occurence of a in φ by x.

4. If φ is a formula, b a free variable of world sort and v a world variable not occuring in φ, then the expressions $\forall v \varphi'$ and $\exists v \varphi'$ are formulas, where φ' is obtained from φ by replacing

each occurence of b in φ by v.

Definition 4.8. An $\mathcal{SL}$**-model** is a triple $\mathfrak{M} = (W, D, V)$, where

1. W and D are non-empty disjoint sets

2. V is a function mapping each $(n+1)$-place predicate symbol P' to a subset $V(P') \subset W \times D^n$.

Definition 4.9. An **assignment** in an $\mathcal{SL}$-model $\mathfrak{M}$ is a function $\mathfrak{a} = \mathfrak{a}_1 \cup \mathfrak{a}_2$, where $\mathfrak{a}_1$ maps every domain variable x to an element $\mathfrak{a}_1(x) \in D$ and $\mathfrak{a}_2$ maps every world variable u to an element $\mathfrak{a}_2(u) \in W$.

The satisfaction relation $\mathfrak{M} \models \varphi[\mathfrak{a}]$ for two-sorted logic is defined in the usual way:

- $\mathfrak{M} \models P'(u, x_1, \ldots, x_n)[\mathfrak{a}]$ iff $(\mathfrak{a}_2(u), \mathfrak{a}_1(x_1), \ldots, \mathfrak{a}_1(x_n)) \in V(P')$,

- $\mathfrak{M} \models \forall u \varphi[\mathfrak{a}]$ iff $\mathfrak{M} = \varphi[\mathfrak{b}]$ for every assignment $\mathfrak{b}$ that differs from $\mathfrak{a}$ at most on v,

- $\mathfrak{M} \models \forall x \varphi[\mathfrak{a}]$ iff $\mathfrak{M} = \varphi[\mathfrak{b}]$ for every assignment $\mathfrak{b}$ that differs from $\mathfrak{a}$ at most on x,

and the standard satisfaction definitions apply to the booleans.

4.3.2 Some properties of two-sorted logic

We have introduced two-sorted logic here in this thesis because we want to translate modal formulas to $\mathcal{SL}$ to talk about their properties there. The tools we can use in two-sorted logic are those that we know from classical first-order logic. We can apply Gentzen's sequent calculus to two-sorted logic just the way we applied it to classical first-order logic in Chapter 2. We only have to pay attention when introducing quantifiers that we can only introduce quantifiers for a variable to replace a term of the same sort, i.e. world or object. The inference rules, however, remain unchanged in principle, meaning that we manipulate the initial sequents with the same rules. Thus, the cut-elimination theorem, Maehara's Lemma and Craig's Interpolation Theorem all apply to two-sorted logic. Consequently, formulas in two-sorted logic interpolate and the translation of modal formulas to two-sorted logic turns out to be a quite useful tool to examine the former.

4.3.3 The standard translation

We can embed the language of quantified modal logic $\mathcal{QML}$ into the two-sorted language $\mathcal{SL}$ in a very straight forward manner. We have already defined the two-sorted predicate symbols in analogy to modal predicate symbols, such that a modal predicate symbol P is associated with the two-sorted predicate symbol P'. Given a world variable v, we define the standard translation ST_v from $\mathcal{QML}$ into $\mathcal{SL}$ this way:

$$ST_v(P(x_1, \dots, x_n)) = P'(v, x_1, \dots, x_n)$$
$$ST_v(\top) = \top$$
$$ST_v(\varphi \wedge \psi) = ST_v(\varphi) \wedge ST_v(\psi)$$
$$ST_v(\neg\varphi) = \neg ST_v(\varphi)$$
$$ST_v(\forall x\varphi) = \forall x ST_v(\varphi)$$
$$ST_v(\Box\varphi) = \forall w ST_w(\varphi)$$

where w is some fresh world variable. Usually the two-sorted language and standard translation also include and consider a predicate symbol R for the accessibility relation R on the worlds W. Since we will only translate formulas in the logic $\mathcal{S5B}$ in this thesis, where we have that $R = W \times W$ for every model, we can omit R. As all worlds are connected to all worlds and we work with constant domains, it consequently does not matter from which world variable v we start the translation ST_v if every subformula is in the scope of some modal operator.

Definition 4.10. Let $\mathfrak{M} = (W, D, V)$ be some $\mathcal{SL}$-model and $\mathfrak{a} = \mathfrak{a}_1 \cup \mathfrak{a}_2$ an assignment in $\mathfrak{M}$. Further, let $\mathfrak{N} = (W', R, D', V')$ be an $\mathcal{S5B}$-model and $\mathfrak{b}$ be an assignment in $\mathfrak{N}$ such that $W = W'$, $D = D'$, $\mathfrak{b} = \mathfrak{a}_1$ and such that for any modal predicate symbol P and counterpart P' in the two-sorted language we have that $V(P) = V'(P')$. Then we call $(\mathfrak{M}, \mathfrak{a})$ and $(\mathfrak{N}, \mathfrak{b})$ **equivalent**, writing $(\mathfrak{M}, \mathfrak{a}) \sim (\mathfrak{N}, \mathfrak{b})$.

Note that for each tuple $(\mathfrak{M}, \mathfrak{a})$ we can find a unique equivalent tuple $(\mathfrak{N}, \mathfrak{b})$. Now that we have defined the standard translation, the question arises whether the translation of a formula that is valid in constant-domain $\mathcal{S5}$ is still valid in $\mathcal{SL}$.

Lemma 4.11. *Suppose we have two equivalent models* $(\mathfrak{M}, \mathfrak{a}) \sim (\mathfrak{N}, \mathfrak{b})$*. For every formula* φ *in quantified modal logic and world variable* v *we then have that*

$$(\mathfrak{N}, \mathfrak{a}(v)) \models \varphi[\mathfrak{b}] \textit{ iff } \mathfrak{M} \models ST_v(\varphi)[\mathfrak{a}].$$

Proof. By induction on the structure of the formula φ. □

Consequently the standard translation is validity preserving. We can also map the $\mathcal{S5B}$-models to models in two-sorted logic bijectively such that we can use them interchangeably: $\mathfrak{M} = (W, R, D, V) \mapsto \mathfrak{N} = (W, D, V)$.

We can argue as well that valid $\mathcal{S5B}$-sentences are still valid when translated to $\mathcal{SL}$ by showing that the translations of all $\mathcal{S5B}$-axioms and inference rules from the Hilbert-style proof system (Section 4.2.2) can be derived in $\mathcal{SL}$ simply using Gentzen's sequent calculus in $\mathcal{SL}$. For example, we can derive the translation of the axiom $\Box A \supset \Box\Box A$, which has the form $\forall v A'(v) \supset \forall v \forall w A'(w)$, the following way:

$$\dfrac{\dfrac{\dfrac{A'(t) \to A'(t)}{\forall v A'(v) \to A'(t)}\ \forall : left}{\forall v A'(v) \to \forall w A'(w)}\ \forall : right}{\forall v A'(v) \to \forall v \forall w A'(w)}\ \forall : right$$

Conveniently, we can easily identify formulas in $\mathcal{SL}$ that are the translation of a modal formula. Looking at the shape of translated formulas, we can note that every subformula is always bound by the innermost world quantifier that it is in the scope of. For example, if we have a formula like $\Box\Diamond P$, then in the translation $\forall w \exists u P'(u)$ the variable u in the predicate P' is bound by the innermost world quantifier $\exists u$. That is, we will not get a formula from the translation in which different world-quantifiers bind the same subformula "cross-wise". The subformula $Q'(v) \wedge P'(w)$ in the formula $\forall w \forall v(Q'(v) \wedge P'(w))$, for example, is in the scope of both world quantifications $\forall w$ and $\forall v$, even though the subformulas $Q'(v)$ and $P'(w)$ are only bound by one of them respectively, and the innermost wold quantifier that the subformula $P'(w)$ is in the scope of does not bind it. In this case, we say that the world-quantifiers **cross-bind**. In contrast, the world quantifiers in the translation $(\forall v(Q'(v) \wedge \exists w P'(w))) \wedge T'(u)$ of the formula $(\Box(Q \wedge \Diamond P)) \wedge T$ do not cross-bind as the subformulas that are being bound are bound by the innermost world-quantifier that they are in the scope of. Thus, formulas in $\mathcal{SL}$ in which world-quantifiers cross-bind express relations between worlds which cannot be expressed in modal logic. Coming from these reflections, we define the *one-world-variable fragment*.

Definition 4.12. Let v be some world variable. Then the **one-world-variable fragment** $\mathcal{SL}^v$ contains the formulas in the two-sorted logic $\mathcal{SL}$ in which every subformula is only bound by the innermost world quantifier. The only free world variable that might appear in a subformula is v, in which case the subformula is not in the scope of any other world-quantifier.

Lemma 4.13. *For every $\mathcal{S5B}$-formula φ, its translation $ST_v(\varphi)$ belongs to the one-world-variable fragment $\mathcal{SL}^v$. Conversely, every formula in such a fragment is the translation $ST_v(\varphi)$ of some formula φ in $\mathcal{QML}$ for some world variable v.*

Proof. The first implication of the lemma can be easily derived from the definition of the translation ST_v. Conversely, we can construct a "re-translation" of formulas in the one-world-variable fragment $\mathcal{SL}^v$ to quantified modal formulas:

$$\begin{aligned} Re_v :\ \mathcal{SL}^v &\longrightarrow \mathcal{QML} \\ P'(v, x_1, \ldots, x_n) &\mapsto P(x_1, \ldots, x_n) \\ \top &\mapsto \top \\ \varphi \wedge \psi &\mapsto Re_v(\varphi) \wedge Re_v(\psi) \\ \neg\varphi &\mapsto \neg Re_v(\varphi) \\ \forall x\varphi &\mapsto \forall x\, Re_v(\varphi) \\ \forall w\varphi &\mapsto \Box Re_w(\varphi) \end{aligned}$$

where w can be any world variable. It is easily checked that for every formula φ in $\mathcal{SL}^v$, we have that $\varphi = ST_v(Re_v(\varphi))$ and thus it is the translation of the formula $Re_v(\varphi)$ in $\mathcal{QML}$. □

The language of quantified modal logic as a tool to talk about models in constant-domain $\mathcal{S}5$ has the same expressive power as the one-variable-fragment $\mathcal{SL}^v$ of two-sorted logic. When we look at what the standard translation does, it really is a tool to translate the semantics of quantified modal logic to syntactic arguments. Whatever we can express semantically in quantified modal logic, we can express syntactically in the translation.

5 Fine's counterexample

We have already discussed that there is no analytic sequent calculus known for the modal logic $\mathcal{S5}$ as we cannot eliminate the cut-rule. Thus, we cannot derive Craig's Interpolation Theorem and consequently Beth's Definability Theorem for quantified $\mathcal{S5B}$ in an analogous manner as in classical first-order logic. In this chapter we will present Kit Fine's result that the Interpolation Theorem does not hold for quantified $\mathcal{S5}$ as demonstrated in his paper "Failures of the Interpolation Lemma in Quantified Modal Logic" (1979) [Fin79]. Fine proceeds by showing that there are formulas in $\mathcal{S5}$ as well as in $\mathcal{S5B}$ with the constant domain axiom scheme $\forall x \Box \varphi \equiv \Box \forall x \varphi$ for which Beth's Definability Theorem does not hold. Consequently, by implication, the Interpolation Theorem does not hold for these logics. As we restrict ourselves to modal logics with constant domains in this thesis, we will concentrate on Fine's counterexample for $\mathcal{S5B}$ in this section.

5.1 Preliminary definitions

To construct a counterexample to Beth's Definability Theorem for $\mathcal{S5B}$, we first need to formulate it for the modal context. There are various ways to formulate implicit and explicit definability in modal logic. We will stick with a formulation that does not involve any modal operators.

Theorem 5.1 (Beth's Definability Theorem for Modal Logic)**.** *Let T be a theory of the logic $\boldsymbol{L}$ and let T' be the result of replacing each occurence of the n-place predicate P in T by a new n-place predicate P'.*
*We will say that P is **implicitly definable** in T if $T, T' \vdash \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \supset P'(x_1, \ldots, x_n))$ in $\boldsymbol{L}$. We say that P is **explicitly definable** in T if there is an L-formula φ in the language of T not containing P such that $T \vdash \forall x_1, \ldots \forall x_n (P(x_1, \ldots, x_n) \equiv \varphi)$ in $\boldsymbol{L}$.*

Then, Beth's Definability Theorem states that the predicate P is implicitly definable in a theory T of the logic $\boldsymbol{L}$ if and only if it is explicitly definable.

We could also define implicit and explicit definability in modal logic by replacing the quantifier prefixes $\forall x_1 \ldots \forall x_n$ with prefixes containing modal operators such as $\Box \forall x_1 \ldots \forall x_n$ or $\Box \forall x_1 \ldots \Box \forall x_n$. According to Fine, Beth's Definability Theorem will still fail for quantified $\mathcal{S5}$ in this case, and he conjectures that it will do the same for quantified $\mathcal{S5}$ with constant domains. We will only consider the version without any modal operators here.

The Interpolation Theorem can be formulated in the general manner: In a logic L, $\vdash \varphi \supset \psi$

implies that $\vdash \varphi \supset \phi$ and $\vdash \phi \supset \psi$ for some formula ϕ in the common language of ϕ and ψ. As shown in Section 3.2, the Interpolation Lemma implies Beth's Definability Theorem.

We will now introduce some notation and definitions that we will need to construct the counterexample.

Definition 5.2. Let $\mathfrak{M} = (W, R, D, V)$ be an $\mathcal{S}5\mathcal{B}$-structure. Then we write $\mathfrak{M}_w$ for the structure (D, V_w), where $V_w(P) = \{(a_1, \ldots, a_n) \in A^n : (w, a_1, \ldots, a_n) \in V(P)\}$. There is a natural notion of isomorphisms for structures of this form. With $\mathfrak{M}_w$ as just defined and structures $\mathfrak{N} = (W', R', D', V')$ and $\mathfrak{N}_u = (D', V'_u)$ we define the isomorphism $\sigma : \mathfrak{M}_w \cong \mathfrak{N}_u$ as:

$$\sigma : \mathfrak{M}_w \to \mathfrak{N}_u$$
$$D \to D'$$
$$V_w \to V'_u$$

where $\sigma : D \to D'$ is a bijective function such that $\sigma : V_w(F) \to V'_u(F)$ is bijective as well for any predicate symbol F in the language.

We will now define an *isomorphism* between two modal structures.

Definition 5.3. For two modal $\mathcal{S}5$-structures $\mathfrak{M} = (W, R, D, V)$ and $\mathfrak{M}' = (W', R', D', V')$, we define an **isomorphism from $\mathfrak{M}$ onto $\mathfrak{M}'$**, written as $\sigma : \mathfrak{M} \cong \mathfrak{M}'$, to be a bijective function $\sigma : D \to D'$ such that

1. $\forall w \in W\ \exists w' \in W' : \sigma : \mathfrak{M}_w \cong \mathfrak{M}'_{w'}$
2. $\forall w' \in W'\ \exists w \in W : \sigma : \mathfrak{M}'_{w'} \cong \mathfrak{M}_w$.

The following lemma will be useful later on.

Lemma 5.4. *Let ρ be an isomorphism $\rho : \mathfrak{M}_w \cong \mathfrak{N}_u$ and let the following condition be fulfilled:*

$$(\forall\, finite\ \rho' \subset \rho)(\exists \sigma \supset \rho')(\sigma : \mathfrak{M} \cong \mathfrak{N}).$$

Then we have that for any formula $\phi(x_1, \ldots, x_n)$ with x_i being free variables, and for $a_1, \ldots, a_n \in D$:

$$(\mathfrak{M}, w) \models \phi(a_1, \ldots, a_n) \text{ iff } (\mathfrak{N}, u) \models \phi(\rho(a_1), \ldots, \rho(a_n)).$$

Proof. By induction. □

5.2 Failures for quantified $\mathcal{S5B}$

We will now construct Fine's counterexample to Beth's Definability Theorem for $\mathcal{S5B}$.

Theorem 5.5. *Beth's Definability Theorem and consequently the Interpolation Theorem do not hold in $\mathcal{S5B}$.*

Proof. We recall that we have the following axioms in $\mathcal{S5B}$:

1. all formulas that have the form of propositional tautologies
2. all instances of the following axiom schemata:
   - $\Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi)$
   - $\Box\phi \supset \phi$
   - $\Box\phi \supset \Box\Box\phi$
   - $\Diamond\Box\phi \supset \phi$
   - $\forall x\Box\phi \equiv \Box\forall x\phi$

Now let T be the theory that we get if we add the following two axioms to $\mathcal{S5B}$:

1. $p \supset \Diamond\forall x(F(x) \supset \Box(p \supset \neg F(x)))$
2. $\neg p \supset \Box\exists(F(x) \wedge \Box(\neg p \supset F(x)))$

We will now show that p is implicitly definable in T, but not explicitly definable. Let us first establish that p is implicitly definable in T:

Let us recall what it means for p to be implicitly definable: If we construct a theory T' by replacing each occurrence of p in T with a new predicate p' of the same degree 0, then p being implicitly definable in T means that the following holds in $\mathcal{S5B}$:

$$T, T' \vdash p \supset p'.$$

Let $\mathfrak{M}$ be a constant domain model (W, R, D, V) and w_0 a world in W such that $\mathfrak{M}$ models T at w_0. For any world $w \in W$, recall that $V_w(F)$ is the subset of D on which F holds in w. Looking at the first axiom that is part of T, we can derive that if $\mathfrak{M}, w_0 \models p$, then there is a world v such that for any element $a \in D$, we have that if $a \in V_v(F)$, then $a \notin V_{w_0}(F)$. Thus, $V_v(F)$ is disjoint from $V_{w_0}(F)$. If we have $\mathfrak{M}, w_0 \not\models p$ on the contrary, then the second axiom tells us that for every world $v \in W$ there is an element a such that $F(a)$ holds in world v and

w_0 alike. Consequently, $V_{w_0}(F)$ is not disjoint from $V_v(F)$ for any world v. Thus we conclude that $\mathfrak{M}, w_0 \models p$ if and only if $V_{w_0}(F)$ is disjoint from $V_v(F)$ for some world $v \in W$. Thus, the extension of p is fixed uniquely in T by the extension of the other symbols in the language. In this way we infer that p is implicitly defined in T.

Next, we will show that p is *not* explicitly definable in T. Remember that p is explicitly definable in T if there is an $\mathcal{S5B}$-sentence φ in the language of T not containing p such that $T \vdash p \equiv \varphi$ in $\mathcal{S5B}$.

To establish the counterexample, we construct an $\mathcal{S5B}$-structure $\mathfrak{U} = (W, D, R, V)$ for the language which has F as its only non-logical predicate. The domain D shall be the set of integers $\mathbb{Z}$. We call a permutation τ on D *essentially finite* if it only permutes finitely many integers and leaves the rest unchanged. We define the set of worlds W as $W = \{\langle k, \tau \rangle : k \in \{0, 1, 2\}$ and τ an essentially finite permutation$\}$. Let $\mathbb{N}$, $\mathbb{O}$ and $\mathbb{E}$ be the set of natural, odd natural and even natural numbers respectively. We set the valuation functions $\{V_w\}_{w \in W}$ to be defined in the following way:

$$V_{\langle 0,\tau \rangle} = \tau(\mathbb{N}) \qquad V_{\langle 1,\tau \rangle} = \tau(\mathbb{O}) \qquad V_{\langle 2,\tau \rangle} = \tau(\mathbb{E}).$$

We write id for the identity permutation on D and denote the world $\langle k, id \rangle$ as w_k for $k = 0, 1$ or 2. Now take ρ to be a permutation on D such that its image $\rho(\mathbb{N}) = \mathbb{O}$. Then we have that $\rho : \mathfrak{U}_{w_0} \cong \mathfrak{U}_{w_1}$. Clearly, $\rho :\ D \to D$ is a bijection on the domains of the two structures. For the valuation of the predicate symbol F we have that

$$\rho :\ V_{\langle 0,id \rangle}(F) = id(\mathbb{N}) = \mathbb{N} \longrightarrow V_{\langle 1,id \rangle}(F) = id(\mathbb{O}) = \mathbb{O} \tag{1}$$

$$n \ \mapsto \ \rho(n) \tag{2}$$

Furthermore, we can establish that $(\forall\ finite\ \rho' \subset \rho)(\exists \sigma \supset \rho')(\sigma : \mathfrak{U} \cong \mathfrak{U})$. Let $\rho' \subset \rho$ be permutation that only acts on a finite set of integers. Take σ to be an essentially finite permutation such that $\sigma \supset \rho'$. Then for every world $\langle k, \tau \rangle$ we have that $\sigma : \mathfrak{U}_{\langle k,\tau \rangle} \cong \mathfrak{U}_{\langle k,\sigma \circ \tau \rangle}$ and conversely $\sigma : \mathfrak{U}_{\langle k,\sigma^{-1} \circ \tau \rangle} \cong \mathfrak{U}_{\langle k,\tau \rangle}$. Thus, by Definition 5.3 of isomorphisms of structures, the function σ is an isomorphism $\sigma : \mathfrak{U} \cong \mathfrak{U}$. Furthermore, we have thus established the conditions of Lemma 5.4 such that we can conclude that for every formula $\phi(x_1, \ldots, x_n)$ with the x_i being free variables, and for $a_1, \ldots, a_n \in D$, we have that:

$$(\mathfrak{U}, w_0) \models \phi(a_1, \ldots, a_n) \text{ iff } (\mathfrak{U}, w_1) \models \phi(\rho(a_1), \ldots, \rho(a_n)).$$

From this we can deduce that

$$(\mathfrak{U}, w_0) \models \varphi \text{ iff } (\mathfrak{U}, w_1) \models \varphi$$

for any closed formula φ.

Now we expand the structure $\mathfrak{U}$ to $\mathfrak{M}$ and $\mathfrak{N}$ by expanding the language to $\{F, p\}$ such that the following holds:

$$(\mathfrak{M}, u) \models p \text{ iff } u \in W \setminus \{w_0\}$$
$$(\mathfrak{N}, u) \models p \text{ iff } u = w_1$$

Consequently, both $(\mathfrak{M}, w_0)$ and $(\mathfrak{M}, w_1)$ model the theory T: The structure $\mathfrak{M}$ models $\neg p$ in the world w_0 and thus automatically models the first axiom of the theory. As we have shown in the first part of the proof, $(\mathfrak{M}, w_0) \models \neg p$ means that $V_{w_0}(F) = \mathbb{N}$ intersects $V_u(F)$ for any world $u \in W$. Since we are dealing with essentially finite permutations, this is clearly true, and thus the second axiom of the theory is fulfilled as well. On the other hand, $(\mathfrak{M}, w_1)$ clearly satisfies the second axiom, and by the above part of the proof $V_{w_1}(F) = \mathbb{O}$ is disjoint from $V_u(F)$ for some $u \in W$. This is the case for $u = w_2$, as $V_{w_2}(F) = \mathbb{E}$. Thus, $(\mathfrak{M}, w_1)$ satisfies the first axiom as well. Now suppose that p was explicitly definable in T, i.e. that there is a sentence θ with sole non-logical constant F such that $T \vdash p \equiv \theta$. Since $(\mathfrak{M}, w_1)$ models T and $(\mathfrak{M}, w_1) \models p$, we infer that $(\mathfrak{M}, w_1) \models \theta$ as well. Therefore, we also have that $(\mathfrak{U}, w_1) \models \theta$. By the consequence of Lemma 5.4 deduced above, we also have that $(\mathfrak{U}, w_0) \models \theta$. This, however, implies that $(\mathfrak{M}, w_0) \models \theta$ and consequently, as $\mathfrak{M}$ is also a model of T at w_0, we get that $(\mathfrak{M}, w_0) \models p$. This is a contradiction of our assumptions about $\mathfrak{M}$. Thus, we have shown that the predicate p is implicitly definable in the theory T, but not explicitly. □

Thus we have seen that Craig's Interpolation Theorem does not hold for all formulas in $\mathcal{S5B}$. However, there are certain classes of formulas in $\mathcal{S5B}$ for which we can derive an interpolant. We will present these in the next chapter.

6 Interpolation by translation in $\mathcal{S5B}$

We have seen in the previous chapter that Craig's Interpolation Theorem does not hold for the modal logic $\mathcal{S5B}$ in general. However, there are some classes of formulas for which it does hold. We will show in this chapter that we can find interpolants for sequents in $\mathcal{S5B}$ consisting of prenex formulas or containing weak modal operators only. Before we start, we will go through some preliminary definitions and results.

6.1 Preliminary definitions and results

Definition 6.1. We define a *prenex formula* in quantified modal logic $\mathcal{QML}$ to be a formula of the form $*A$ where $*$ is an arbitrary combination of quantifiers and modal operators, and A is free from both.

We call a sequent a *prenex sequent* if it consists of prenex formulas only.

Definition 6.2. Let $\Gamma \to \Delta$ be a sequent in $\mathcal{QML}$. Then we call the modal operators $\Diamond$ and $\Box$ appearing in the sequent *strong* (or conversely *weak*) modal operators if in the translation of the sequent the respective quantifiers would be considered *strong* (or *weak*).

Lemma 6.3. *Let $*A$ be a prenex formula in $\mathcal{S5B}$ containing at least one modal operator. Then only the innermost modal operator has a binding effect on A, such that all other modal operators can be discarded while maintaining equivalence.*

Proof. We will expand on our argument from introducing the one-world-variable fragment in Section 4.3.3. Let $*A$ be a prenex formula in $\mathcal{S5B}$. We know that in modal logics, predicates may have different extensions in every world. Since we look at prenex formulas here, we only interpret predicate symbols in the worlds pointed to by the innermost modal operator. Using Kripke semantics, the world relation R in $\mathcal{S5B}$ connects all worlds, that is, $R = W \times W$. Consequently, the worlds pointed to by the innermost modal operator are either *all* worlds if the operator is $\Box$, or *any* world if the operator is $\Diamond$. Thus we interpret the scope of the innermost modal operator independently of the modal operators appearing before it in the formula's prefix. We will refer to these uninterpreted modal operators as outer modal operators.

Since we have constant domains for the scope of the individual variables, their position within the prefix $*$ relative to the outer modal operators is irrelevant as well, as their scope is the same in every world. Consequently, if we discard of the modal operators except for the innermost, we receive a logically equivalent formula, as these operators are not relevant for the interpretation of the formula in $\mathcal{S5B}$.

This property can also be nicely illustrated by translating the formula using the standard translation to two-sorted logic. Let us take the formula $\varphi = \Box\exists x\Diamond\forall y Q(x, y)$ for example. The standard translation of this formula is $\forall v\exists x\exists w\forall y Q'(w, x, y)$. We see that the predicate symbol Q' is not bound by the outer wold quantifier but only depends on the individual quantified variables and the innermost world quantification. □

6.2 Interpolation of prenex formulas in $\mathcal{S5B}$

In this section we will show a way to find the interpolant for sequents consisting of two prenex formulas in $\mathcal{S5B}$. The idea is to translate the sequent to two-sorted logic, find the interpolant there, and translate it back to $\mathcal{S5B}$. Since the prenex formula structure is preserved by the translation, we can apply the Midsequent Theorem (Theorem 3.4) and thus have a proof of the translated sequent at hand that makes it easy to infer an interpolant that can be re-translated. The procedure follows these steps:

1. Translate the prenex formula to $\mathcal{SL}$
2. Apply the Midsequent Theorem
3. Find the (propositional) interpolant of the midsequent
4. Use Maehara's Lemma to construct the interpolant of the translated formula from the midsequent's interpolant
5. Translate the interpolant back to $\mathcal{QML}$

This procedure works for prenex formulas because the interpolant derived by Maehara's Lemma in two-sorted logic will be in prenex form and in the one-variable fragment $\mathcal{SL}^v$ for some world variable v as well. This way, it can be easily translated back to $\mathcal{S5B}$.

Theorem 6.4. *Let $\varphi \supset \psi$ be a valid implication in $\mathcal{S5B}$, where both the antecedent and succedent are prenex formulas of the form $\varphi = *_A A$ and $\psi = *_B B$ and both prefixes contain at least one modal operator. Then the implication has an interpolant γ in $\mathcal{S5B}$. That is, there is a formula γ that is in the common language of φ and ψ, or that consists of $\top$ and logical symbols only, such that $\varphi \supset \gamma$ and $\gamma \supset \psi$ are valid in $\mathcal{S5B}$.*

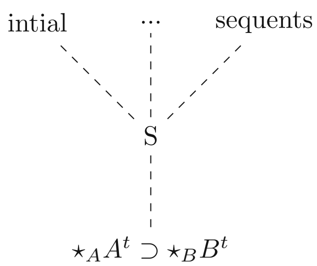
Proof. Let $\varphi \supset \psi$ be a valid implication in $\mathcal{S5B}$, where both the antecedent and succedent are prenex formulas of the form $\varphi = *_A A$ and $\psi = *_B B$. Since we work in $\mathcal{S5B}$, only the innermost modal operator has a binding effect on the formulas A and B by Lemma 6.3, such that we can assume without loss of generality that there is at most one modal operator present in the prefixes $*_A$ and $*_B$ respectively. We will proceed in this proof by first translating the formula to $\mathcal{SL}$, applying the Midsequent Theorem and finding its interpolant there using Craig's Interpolation Theorem. We will deduce that this interpolant is in prenex form from the midsequent structure of the proof. We will continue by showing that the formulas in the antecedent and succedent of the midsequent are already in a one-world-variable fragment. We will use this to show that the interpolant must be in a one-world-variable fragment as well. This way we can translate it back to $\mathcal{S5B}$ and receive an interpolant for the original formula.

We translate the given implication to a formula in two-sorted logic:

$$\begin{aligned} ST_w(\varphi \supset \psi) &= ST_w(*_A A \supset *_B B) \\ &= \star_A ST_v(A) \supset \star_B ST_v(B) \\ &= \star_A A^t \supset \star_B B^t \end{aligned}$$

where the quantifier prefixes $\star_A$ and $\star_B$ result from the translation of the prefixes $*_A$ and $*_B$ respectively. The formulas A^t and B^t stand for $ST_v(A)$ and $ST_v(B)$. Since A and B are free from modal operators as well as quantifiers, their translations $ST_v(A)$ and $ST_v(B)$ are also quantifier-free. Thus, the translation is an implication of prenex formulas as well.

By Craig's Interpolation Theorem (Theorem 3.6), the implication $\star_A A^t \supset \star_B B^t$ has an interpolant $\tilde{C}$. To deduce the interpolant, we apply the Midsequent Theorem (Theorem 3.4) to the sequent $\star_A A^t \supset \star_B B^t$ to obtain a cut-free proof thereof of the following form:



where S is the midsequent of the proof. According to the Midsequent Theorem, every inference above the sequent S is either structural or propositional, and every inference below is either a structural or a quantifier inference. Thus, the midsequent has a propositional, i.e. quantifier-free interpolant C. By applying the steps of the proof of Maehara's Lemma (Lemma 3.5) inductively, we can construct the interpolant $\tilde{C}$ step by step from C. From these steps we see that each structural inference as well as the introduction of strong quantifiers leaves the interpolant unchanged. When introducing a weak quantifier though, the interpolant is also quantified in the way described in the proof of Maehara's Lemma. As the only way that the interpolant C is modified in these steps is by adding quantifier prefixes, we know that the interpolant $\tilde{C}$ is in prenex form as well.

It remains to show that $\tilde{C}$ is in a one-variable fragment for a variable of sort world. We will argue that before any quantifiers are introduced to the interpolant, it contains at most one world variable, such that upon the introduction of a world quantifier, the result is in a one-world-variable fragment as well. We will deduce this from the fact that the formulas in the antecedent and succedent are in a one-world-variable fragment as well.

We already know that the translated endsequent $\star_A A^t \supset \star_B B^t$ is in a one-world-variable fragment. We assumed without loss of generality that both the formula in the antecedent and the one in the succedent contain exactly one world-quantification as we only need to consider the innermost world quantifier. Looking at the proof, when quantifier inferences are applied to a formula to introduce a world quantifier, then this formula can contain only the one free world variable that is being replaced by the bound world variable. Otherwise we would get a formula in the endsequent that is not in a one-world-variable fragment anymore and in which world-quantifiers cross-bind, as the world variables in the formula would not be bound by the innermost world quantifier that they are in the scope of. For example, a formula in the succedent containing more than one free world variable could look like $A'(u, x) \wedge B'(w, y)$ before the quantifier inference. Upon quantification, we would have to introduce the same world variable for both u and w for the world quantifiers not to cross-bind and to end up in a one-world-variable fragment. We would get, for example, the expression $\exists v \exists v(A'(v, x) \wedge B'(v, y))$. However, this is not allowed as a formula by definition. On the other hand, introducing two different quantifiers such as $\exists w \exists v(A'(v, x) \wedge B'(w, y))$ would get us a formula that contains more than one world-variable quantification and that is not in a one-world-variable fragment anymore. As the proof is cut-free, this contradicts the endsequent being in a one-world-variable fragment. Thus, we have only one free world variable within the affected formula upon introduction of world quantifiers. Consequently, the formulas in the antecedent and succedent of the midsequent are in a one-world-variable fragment.

We can inductively deduce that the interpolant is in the one-world-variable fragment as well.

As we start our proof from atomic sequents of the form $A \rightarrow A$ and $\rightarrow \top$, the interpolant starts off containing only the free world variable that is present in the antecedent and succedent of the atomic formulas from the start (if any). Further, once the interpolant contains a free world variable, we cannot add any new world variable to it by any inference as they either do not affect the interpolant or would create formulas in the endsequent that are not in a one-world variable fragment. For example, if a propositional rule like $\wedge: left$ introduces a new subformula with a world variable when the affected formula already contains a world variable, it does not affect the interpolant and it would further contradict the formulas in the antecedent and succedent being in a one-world-variable fragment in the midsequent already. If in the rule $\vee: left$ two formulas with different world variables are being combined, we would create formulas in the midsequent that are not in the one-world-variable fragment as well. Thus the only way to introduce a new world variable to the antecedent or succedent is through a weakening. This leaves the interpolant unchanged. Thus, if the antecedent and succedent of the midsequent are logically connected, i.e. the interpolant is not only made up of $\top$ and logical symbols only, then the interpolant contains only one free world-variable that the antecedent and succedent have in common.

Now we know that if the interpolant C of the midsequent does not consist of $\top$ and logical symbols only, it contains only one kind of free world variable. Any world quantifier introduction on the interpolant will bind this free variable by the proof of Maehara's Lemma, as there are no quantifier introductions on interpolants that do not actually bind any variable. Thus, the final interpolant of the endsequent is in the one-variable fragment $\mathcal{SL}^w$ as well for some world variable w. Consequently, it is re-translatable to $\mathcal{QML}$ and by the truth-maintenance of the translation, it is an interpolant of our original formula $*_A A \supset *_B B$.

□

Using this translation procedure, we can also find the interpolant for valid implications in $\mathcal{S5B}$ that contain weak modal operators only:

Theorem 6.5. *Let $\phi \supset \psi$ be a valid implication in $\mathcal{S5B}$, where both the antecedent and succedent contain arbitrary quantifiers but only weak modalities. Then the implication has an interpolant γ in $\mathcal{S5B}$. That is, there is a formula γ that is in the common language of φ and ψ, or that consists of $\top$ and logical symbols only, such that $\varphi \supset \gamma$ and $\gamma \supset \psi$ are valid in $\mathcal{S5B}$.*

Proof. Let $\varphi \supset \psi$ be a valid implication in $\mathcal{S5B}$, where both the antecedent and succedent contain arbitrary quantifiers but only weak modalities. Let $\varphi' \supset \psi'$ be the formula's translation to $\mathcal{SL}$, P a cut-free proof thereof, and γ' its interpolant according to Craig's Interpolation Lemma, i.e. the interpolant derived using the construction steps of the proof of Maehara's Lemma. We want to find out whether we can derive the interpolant of the original modal formula from this one. Since we only had weak modalities in $\varphi \supset \psi$ in the first place, we only have weak world quantifiers in the translation $\varphi' \supset \psi'$. Consequently, we do not have any eigenvariable conditions in the inferences introducing the world variables in the proof P. Further, we know that the endsequent is a translation of a modal formula and thus in a one-world-variable fragment. Analogously to the argumentation in the proof of Theorem 6.4, when a world quantifier is introduced for a formula in the course of the proof, this formula only has at most a single free world variable before that, in which case it is being bound by the quantifier introduction. Else we would get cross-binding quantifiers and not end up with a translation of a modal formula in the end-sequent as our proof is cut-free.

With this in mind, we can set all free world variables to the same variable in the proof. Consequently, the interpolant γ' has only one free world variable as well in the course of the proof, and every introduction of world quantifiers then binds this variable in the respective formula. Thus, we have that γ' is in a one-world-variable fragment and a translation of an $\mathcal{S5B}$-formula γ. By the truth maintaining properties of the translation, γ is an interpolant for $\varphi \supset \psi$. □

We can also show in an analogous manner that propositional $\mathcal{S5}$ interpolates.

Corollary 6.5.1. *Propositional $S5$ interpolates.*

Proof. Let $\varphi \supset \psi$ be a valid implication in propositional $\mathcal{S5}$. By a simple embedding, we can interpret the implication as a formula in $\mathcal{S5B}$ without any quantifiers and with 0-place predicates only. Translating $\varphi \supset \psi$ to $\mathcal{SL}$, we get a formula $\varphi' \supset \psi'$ that consists of 1-place predicates only, where the arguments are world arguments. Since $\varphi' \supset \psi'$ is the translation of a modal formula, there is at most one kind of free world variable, say v, present in it. Applying Craig's Interpolation Theorem, we receive an interpolant γ' of the implication $\varphi' \supset \psi'$. Since we only have variables from the sort world, we can treat the formula like one in classical first-order logic. Consequently, we can move quantifiers around – analogously to the quantifier movements in order to create a logically equivalent prenex normal form – to create a logically equivalent formula where each quantifier only has predicates in its scope that it actually binds. The modified formula only contains 1-place predicates, thus no predicate is in the scope of more than one quantifier. Thus, we can replace all bound variables with the world variable w and receive a logically equivalent formula. Since this formula belongs to the one-world-variable fragment $\mathcal{SL}^v$, we can translate it back to $\mathcal{S5B}$. The resulting formula γ belongs to propositional modal logic, as we have 0-place predicates again. Since the translation maintains validity, the formula γ is an interpolant of $\varphi \supset \psi$. $\square$

Corollary 6.5.2. *Let $\varphi \to \tilde{C}$ be a valid prenex sequent in $\mathcal{S5B}$ of the form $\varphi \to *_C \Box C$, where the prefix $*_C$ stands for an arbitrary combination of existential quantifiers and modal operators, $\Box$ is the innermost modality of the succedent, and C is some formula without any modalities but possibly including more quantifiers. Then we can constructively derive unique terms that serve as truth witnesses for the quantifiers in $*_C$.*

Proof. Let $\varphi \to *_C \Box C$ be a valid sequent in $\mathcal{S5B}$ of the form given in the corollary. We translate this modal formula to $\mathcal{SL}$ and get the sequent

$$\varphi' \to *_{C'} \forall v C'(v),$$

where φ' is the translation of the formula φ, $\forall v C'(v)$ is the translation of $\Box C$, and the quantifier prefix $*_{C'}$ results from the translation of the prefix $*_C$.

We can argue analogously to the proof of Theorem 6.4 that the translated sequent consists of prenex formulas only, and that there is a cut-free proof in the form given by the Midsequent Theorem. Similarly we can assume that due to the requirement that the endsequent is in the one-world-variable fragment, the formulas in the antecedent and the succedent of the midsequent are as well. We also deduced in the proof of Theorem 6.4 that weakenings that are introduced in order to enable propositional inferences above the midsequent do not introduce any new world variable to the proof. Else the formula resulting from the propositional connection would contain more than one free world variable. Since the antecedent and succedent of the endsequent consist of only one formula each, we do not need any weakenings that would introduce additional world variables, unless to prove a sequent in which antecedent and succedent are not logically connected, such as for example the sequent $\exists x \Box B(x) \to \exists x \Box (A(x) \wedge \neg A(x))$. In this case a single weakening suffices to add the missing formula.

Thus, at the point of the proof when the $\forall v$ quantification is introduced to the succedent, it does not necessarily contain two copies with different free world variables each of the formula that is being quantified. For example, let the quantification rule be applied to the formula $D(u, s_1, \ldots, s_n)$, where u is a free world variable and $s_1, \ldots s_n$ are the terms of domain sort that are being existentially quantified over in the endresult. Then the succedent does not have to contain a copy $D(v, s_1, \ldots, s_n)$ with a different world variable, as this would have been introduced by a weakening which we do not need for any propositional inference. The succedent does not contain a copy $D(u, s_1, \ldots, s_n)$ with the same world variable either, as this would violate the eigenvariable condition. Thus there is only one copy of $D(u, s_1, \ldots, s_n)$. Consequently, D is not being contracted after this quantifier introduction, and $s_1, \ldots, s_n$ are unique terms that serve as witnesses for the existential quantifiers outside of $\forall v$ in ψ'.

Since the succedent of the endsequent $*_{C'} \forall v C'(v)$ is still in a one-world-variable fragment if we strip it off the outer existential quantifiers and replace the respective variables by the place holders before their quantifier introductions, we can translate the modified formula back to $\mathcal{S5B}$ and get the desired result.

□

6.3 Examples

We will begin with an example to illustrate the statement of Theorem 6.4 that we can derive the interpolant of sequents of prenex formulas in $\mathcal{S5B}$ by translation to $\mathcal{SL}$.

Example 6.6. We will show by translation to $\mathcal{SL}$ that the $\mathcal{S5B}$-formula $\Box\forall x(A(x) \wedge B) \to \Box\exists x(A(x) \vee C)$ interpolates.

1. Translate the prenex formula to $\mathcal{SL}$:

$$\forall v \forall x(A'(v, x) \wedge B'(v)) \to \forall v \exists x(A'(v, x) \vee C'(v))$$

2. Apply the Midsequent Theorem:

   We get the following cut-free proof of the translated sequent that contains the midsequent $M = A'(s, t) \wedge B'(s) \to A'(s, t) \vee C'(s)$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A'(s,t) \to A'(s,t)}{A'(s,t) \wedge B'(s) \to A'(s,t)}\ \wedge : left}{A'(s,t) \wedge B'(s) \to A'(s,t) \vee C'(s)}\ \vee : right}{\forall x(A'(s,x) \wedge B'(s)) \to A'(s,t) \vee C'(s)}\ \forall : left}{\forall x(A'(s,x) \wedge B'(s)) \to \exists x(A'(s,x) \vee C'(s))}\ \exists : right}{\forall v \forall x(A'(v,x) \wedge B'(v)) \to \exists x(A'(s,x) \vee C'(s))}\ \forall : left}{\forall v \forall x(A'(v,x) \wedge B'(v)) \to \forall v \exists x(A'(v,x) \vee C'(v))}\ \forall : right$$

3. Find the (propositional) interpolant of the midsequent:

   The interpolant of the midsequent $A'(s, t) \wedge B'(s) \to A'(s, t) \vee C'(s)$ is the formula $A'(s, t)$.

4. Use Maehara's Lemma to construct the interpolant of the translated formula from the midsequent's interpolant:

   We will apply the steps from the proof of Maehara's Lemma (Lemma 3.5) starting from the midsequent's interpolant to construct the interpolant of the endsequent and will denote the inference rule which we are considering on the left:

$$\begin{array}{rc} \forall : left & \dfrac{A'(s,t)}{\forall x A'(s,x)} \\ \exists : right & \dfrac{}{\forall x A'(s,x)} \\ \forall : left & \dfrac{}{\forall v \forall x A'(v,x)} \\ \forall : right & \dfrac{}{\forall v \forall x A'(v,x)} \end{array}$$

   Thus, we get the interpolant $\forall v \forall x A'(v,x)$ for the translated sequent $\forall v \forall x(A'(v,x) \wedge B'(v)) \to \forall v \exists x(A'(v,x) \vee C'(v))$.

5. Translate the interpolant back to $\mathcal{S5B}$:

   The retranslation of the interpolant $\forall v \forall x A'(v,x)$ yields the interpolant $\Box \forall x A(x)$ for the original sequent $\Box \forall x(A(x) \wedge B) \to \Box \exists x(A(x) \vee C)$.

The method presented above, where we first translate an $\mathcal{S5B}$-formula to two-sorted logic and find the interpolant using Maehara's Lemma, can be applied to non-prenex formulas as well in order to find their interpolant in $\mathcal{SL}$. In this case, however, we cannot apply the Midsequent Theorem to obtain a certain proof structure, as the formula lacks the necessary prenex form. As we will see, in those cases following the steps of Maehara's Lemma does not always result in an interpolant that can be directly translated back to $\mathcal{S5B}$, even though the original modal formula might have an interpolant. We will present an example of this below.

Example 6.7. Let C, D, P and Q be predicate symbols in $\mathcal{S5B}$. We will try to find the interpolant of the following valid sequent containing non-prenex formulas:

$$\Diamond\big((P \vee P) \wedge C\big),\ \Diamond\big((Q \vee Q) \wedge C\big) \to \Diamond(P \vee D) \wedge \Diamond(Q \vee D),\ \Diamond P,\ \Diamond Q$$

As can be easily checked, a suitable interpolant in $\mathcal{S5B}$ would be the formula $\Diamond P \wedge \Diamond Q$. Let us see what kind of interpolant we get from applying Maehara's Lemma. We will switch to a landscape view in order to improve readability.

1. Translate to $\mathcal{SL}$:

$$\exists v\big((P'(v) \lor P'(v)) \land C'(v)\big),\ \exists v\big((Q'(v) \lor Q'(v)) \land C'(v)\big) \ \rightarrow\ \exists v(P'(v) \lor D'(v)) \land \exists v(Q'(v) \lor D'(v)),\ \exists v P'(v),\ \exists v Q'(v)$$

2. Use Maehara's Lemma to construct the interpolant from the proof of the translated formula: We give a proof here in which we do not explicitly mention the *Exchange* rule when used, and the abbreviation W stands for the $Weakening$ rule. We first write both beginning branches of the prooftree separately and combine them later on for space reasons:

$$\lor : left\ \dfrac{P'(u) \rightarrow P'(u) \qquad P'(u) \rightarrow P'(u)}{P'(u) \lor P'(u) \rightarrow P'(u)}$$

$$W : right\ \dfrac{P'(u) \lor P'(u) \rightarrow P'(u)}{P'(u) \lor P'(u) \rightarrow P'(u),\ P'(u)}$$

$$\lor : right\ \dfrac{P'(u) \lor P'(u) \rightarrow P'(u),\ P'(u)}{P'(u) \lor P'(u) \rightarrow P'(u) \lor D'(u),\ P'(u)}$$

$$\land : left\ \dfrac{P'(u) \lor P'(u) \rightarrow P'(u) \lor D'(u),\ P'(u)}{(P'(u) \lor P'(u)) \land C'(u) \rightarrow P'(u) \lor D'(u),\ P'(u)}$$

$$\exists : right\ \dfrac{(P'(u) \lor P'(u)) \land C'(u) \rightarrow P'(u) \lor D'(u),\ P'(u)}{(P'(u) \lor P'(u)) \land C'(u) \rightarrow \exists v(P'(v) \lor D'(v)),\ P'(u)}$$

$$W : left\ \dfrac{(P'(u) \lor P'(u)) \land C'(u) \rightarrow \exists v(P'(v) \lor D'(v)),\ P'(u)}{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(P'(v) \lor D'(v)),\ P'(u)}$$

$$W : right\ \dfrac{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(P'(v) \lor D'(v)),\ P'(u)}{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(P'(v) \lor D'(v)),\ P'(u),\ Q'(b)}$$

$$\lor : left\ \dfrac{Q'(b) \rightarrow Q'(b) \qquad Q'(b) \rightarrow Q'(b)}{Q'(b) \lor Q'(b) \rightarrow Q'(b)}$$

$$W : right\ \dfrac{Q'(b) \lor Q'(b) \rightarrow Q'(b)}{Q'(b) \lor Q'(b) \rightarrow Q'(b),\ Q'(b)}$$

$$\lor : right\ \dfrac{Q'(b) \lor Q'(b) \rightarrow Q'(b),\ Q'(b)}{Q'(b) \lor Q'(b) \rightarrow Q'(b) \lor D'(b),\ Q'(b)}$$

$$\land : left\ \dfrac{Q'(b) \lor Q'(b) \rightarrow Q'(b) \lor D'(b),\ Q'(b)}{(Q'(b) \lor Q'(b)) \land C'(b) \rightarrow Q'(b) \lor D'(b),\ Q'(b)}$$

$$\exists : right\ \dfrac{(Q'(b) \lor Q'(b)) \land C'(b) \rightarrow Q'(b) \lor D'(b),\ Q'(b)}{(Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(Q'(v) \lor D'(v)),\ Q'(b)}$$

$$W : left\ \dfrac{(Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(Q'(v) \lor D'(v)),\ Q'(b)}{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(Q'(v) \lor D'(v)),\ Q'(b)}$$

$$W : right\ \dfrac{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(Q'(v) \lor D'(v)),\ Q'(b)}{(P'(u) \lor P'(u)) \land C'(u),\ (Q'(b) \lor Q'(b)) \land C'(b) \rightarrow \exists v(Q'(v) \lor D'(v)),\ P'(u),\ Q'(b)}$$

We combine the last sequents of the two prooftrees above using the rule $\wedge : right$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\text{Endsequent of the first prooftree} \qquad \text{Endsequent of the second prooftree}}{(P'(u) \vee P'(u)) \wedge C'(u),\ (Q'(b) \vee Q'(b)) \wedge C'(b) \rightarrow \exists v(P'(v) \vee D'(v)) \wedge \exists v(Q'(v) \vee D'(v)),\ P'(u),\ Q'(b)}\ \wedge : right}{(P'(u) \vee P'(u)) \wedge C'(u),\ (Q'(b) \vee Q'(b)) \wedge C'(b) \rightarrow \exists v(P'(v) \vee D'(v)) \wedge \exists v(Q'(v) \vee D'(v)),\ \exists v P'(v),\ Q'(b)}\ \exists : right}{(P'(u) \vee P'(u)) \wedge C'(u),\ (Q'(b) \vee Q'(b)) \wedge C'(b) \rightarrow \exists v(P'(v) \vee D'(v)) \wedge \exists v(Q'(v) \vee D'(v)),\ \exists v P'(v),\ \exists v Q'(v)}\ \exists : right}{\exists v\big((P'(v) \vee P'(v)) \wedge C'(v)\big),\ (Q'(b) \vee Q'(b)) \wedge C'(b) \rightarrow \exists v(P'(v) \vee D'(v)) \wedge \exists v(Q'(v) \vee D'(v)),\ \exists v P'(v),\ \exists v Q'(v)}\ \exists : left}{\exists v\big((P'(v) \vee P'(v)) \wedge C'(v)\big),\ \exists v\big((Q'(v) \vee Q'(v)) \wedge C'(v)\big) \rightarrow \exists v(P'(v) \vee D'(v)) \wedge \exists v(Q'(v) \vee D'(v)),\ \exists v P'(v),\ \exists v Q'(v)}\ \exists : left$$

By inductively reconstructing the interpolant using Maehara's Lemma, we get the interpolant $I' = \exists w \exists v(P'(v) \wedge Q'(w))$.

3. Can we translate the interpolant back to $\mathcal{S5B}$? If we naively translate it back using an analogue of the re-translation defined for formulas in the one-variable fragment $\mathcal{SL}^v$, we get the corresponding formula $\Diamond\Diamond P \wedge Q$, where the first diamond should refer to P and the second one should bind Q, a specification which cannot be expressed in modal logic in this way. As we expressed in 4.13, translations of modal formulas always belong to a one-world-variable fragment $\mathcal{SL}^v$ for some world-variable v, which clearly is not the case here. As we have already established above, another suitable interpolant would be $\Diamond P \wedge \Diamond Q$. The translation of this interpolant is $\exists v P'(v) \wedge \exists v Q'(v)$, which is logically equivalent to I'. Thus, for non-prenex sequents, Maehara's Lemma does not always return an interpolant that is re-translatable, even if there exists an interpolant for the original formula in $\mathcal{S5B}$, as without the prenex form we do not always get an interpolant in a one-world-variable fragment.

We will continue with an example that illustrates that propositional $\mathcal{S}5$ interpolates with the method of Corollary 6.5.1.

Example 6.8. We will just continue with the previous example, as it can be read as a propositional formula.

$$\Diamond\big((P \vee P) \wedge C\big),\ \Diamond\big((Q \vee Q) \wedge C\big) \rightarrow \Diamond(P \vee D) \wedge \Diamond(Q \vee D),\ \Diamond P,\ \Diamond Q$$

The interpolant we received from applying Craig's Interpolation Theorem to the translation of the sequent is $I' = \exists w \exists v (P'(v) \wedge Q'(w))$. By moving the variables while maintaining equivalence to disentangle the formula, we get the interpolant $\exists v P'(v) \wedge \exists w Q'(w)$. We replace all bound variables by the same variable to get the equivalent formula $\exists w P'(w) \wedge \exists w Q'(w)$, which is clearly a translation of the modal formula $\Diamond P \wedge \Diamond Q$ and interpolant of the original sequent.

7 Conclusion

In this thesis, we examined the possibilities and limitations of interpolation in the modal system $\mathcal{S5B}$. Kit Fine showed that implicit and explicit definability is not equivalent in $\mathcal{S5B}$. Consequently, we cannot find an interpolant for every formula in this system. There are exceptions, however. We showed that sequents consisting of prenex formulas or containing arbitrary quantifiers but only weak modal operators allow for interpolation. In a next step one could examine whether the same goes for dual sequents - those with weak quantifiers only but arbitrary modal operators.

When we translate formulas from propositional $\mathcal{S}5$ to the language of two-sorted logic $\mathcal{SL}$, then we can find interpolants there that might or might not be re-translatable to $\mathcal{S}5$. That is, we will presumably find more interpolants for the same sequent in $\mathcal{SL}$ than in $\mathcal{S}5$. For an interpolant that is re-translatable, different world quantifiers that were cross-binding predicates are disentagled such that they only have predicates in their scope that they actually bind. This disentaglement, however, might prolong the interpolant. Interpolants that are not disentagled, where quantifiers might cross-bind predicates, thus making it impossible to re-translate them to $\mathcal{S}5$, could be exponentially smaller than disentagled ones. Thus the size of not re-translatable interpolants might be an interesting topic of further research. Disentangling is an exponential decision procedure. Once disentangled, satisfiability is an NP decision problem, just as satisfiability in propositional logic is.

Another question worth looking at is to which extend proofs might be shorter for the translation of a sequent in $\mathcal{SL}$ than in the original $\mathcal{S}5$, provided we allow the cut rule to be used in $\mathcal{SL}$. We cannot necessarily re-translate the cut formulas to $\mathcal{S}5$. It would be worth examining what effect those cuts have on the length of the proof. The conjecture is that the proofs in $\mathcal{SL}$ are shorter than in $\mathcal{S}5$ as its language is more comprehensive. The difference might be non-elementary (that is, it might grow faster than $2^{\cdot^{\cdot^{2^n}}}$). The worst case complexity of cut-elimination is non-elementary for classical first-order logic ([Mat15], p. 1). This topic invites further examination.

References

[BG07] Torben Braüner and Silvio Ghilardi. Modal Logic: A Semantic Perspective. In Patrick Blackburn et al., editor, *Handbook of Modal Logic*, volume 3, chapter 9. Elsevier, 2007.

[Bur] Samara Burns. Hypersequent Calculi for Modal Logics (Unpublished master's thesis). *University of Calgary, Calgary*, retrieved from: https://prism.ucalgary.ca/handle/1880/106539 on 18.10.2022.

[BvB07] Patrick Blackburn and Johan van Benthem. Modal Logic: A Semantic Perspective. In Patrick Blackburn et al., editor, *Handbook of Modal Logic*, volume 3, chapter 1. Elsevier, 2007.

[Fin79] Kit Fine. Failures of the Interpolation Lemma in Quantified Modal Logic. *The Journal of Symbolic Logic*, 44(2):201–206, 1979.

[Gen35] Gerhard Gentzen. Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 30(6):175–210, 405–431, 1935.

[LR15] Björn Lellmann and Revantha Ramanayake. Lecture 3 - Modal Logic S5 and Hypersequents. In *Proof Theoretical Reasoning Lecture Notes*. retrieved from: https://www.logic.at/staff/lellmann/static/talks/tutorial2015TRS1.pdf on 08.01.2022, 2015.

[Mat15] Matthias Baaz and Christian G. Fermüller. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 94–109, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[OM57] Masao Ohnishi and Kazuo Matsumoto. Gentzen Method in Modal Calculi. *Osaka Math. J.*, (9):113–130, 1957.

[Pfe84] Frank Pfenning. Analytic and Non-analytic Proofs. In *Shostak, R.E. (eds) 7th International Conference on Automated Deduction. CADE 1984. Lecture Notes in Computer Science*, volume 170, page 394–413. Springer, New York, NY, 1984.

[Pri08] Graham Priest. *Einführung in die nicht-klassische Logik.* Cambridge University Press, 2nd edition, 2008.

[Ste88] Peter Steinacker. Modallogik. In Kreiser et al., editor, *Nichtklassische Logik - eine Einführung*, chapter 3. Akademie-Verlag Berlin, 1988.

[SW00] Holger Sturm and Frank Wolter. First-order Expressivity for S5-models: Modal vs. Two-sorted Languages. *Journal of Philosophical Logic*, 44(2):201–206, 2000.

[Swi98] Zeno Swijtink. Beth's theorem and Craig's theorem. In Edward Craig, editor, *The Routledge Encyclopedia of Philosophy*, pages 760–763. Taylor and Francis, 1998.

[Tak87] Gaisi Takeuti. Proof Theory. Studies in Logic and the Foundations of Mathematics. North-Holland, 2nd edition, 1987.