Understanding Definability in First-Order Logic

by

Joël Ravelomanantsoa-Ratsimihah

Submitted to

Central European University

Department of Mathematics and Applications

In partial fulfillment of the requirements for the degree of Master of Science

Supervisor: Professor Ildikó Sain, Senior Research Fellow at Alfréd Rényi Institute of Mathematics

Budapest, Hungary

2010

Table of Contents

Chapter 1

Introduction

Definability theory is a branch of model theory which has various applications in several fields of research, e.g. in *theoretical physics*, *theoretical computer science*, *algebraic logic*.

As it has been pointed out in various works (e.g. [1] and [14]), definability was one of Alfred Tarski's favourite subjects already in the 1930's. In the paper [15] he formulated and started the project of bringing about a definability theory.

The fact that, before exploring logic, Tarski did research in sciences and in the methodology of science indicates that he might well have motivations coming from his scientific experience. In this line it is remarkable that Hans Reichenbach, in his book [13] (already in 1920), explains that definability is a basic factor in *relativity theory*. This idea appears already in Einstein's work, in 1905, but more implicitly than in [13].

Another source of motivation can be the pioneering paper [6], where Willem Blok and Don Pigozzi explain that definability is a corner stone of *algebraic logic*. To illustrate this, let us recall from any textbook on the subject (cf. e.g. [12], [10], [11], [2]) that the algebraic version $\mathfrak{Cs}(\mathfrak{M})$ of a model $\mathfrak{M}$ of first order logic is an algebra the universe of which is the collection

$$\{\varphi^{\mathfrak{M}} : \varphi \text{ a formula}\}$$

of all *definable relations* of the model; where

$$\varphi^{\mathfrak{M}} := \{k \in {}^{\omega}M \text{ such that } \mathfrak{M} \models \varphi[k]\}$$

is the meaning of φ in $\mathfrak{M}$.

The starting point of definability theory is the following. Given a theory, how to determine whether some property r is definable in terms of certain other notions? Suppose that L is a first-order language and L' is the first-order language that we get from L by adding a new predicate symbol r. Suppose also that T is a set of formulas of L'. We have the following definitions.

Definition 1.1. We say that the theory T *defines* r *explicitly* if and only if there is a formula $\varphi(x_1, \ldots, x_n)$ of L such that in every model of T, the formulas $\varphi(x_1, \ldots, x_n)$ and $r(x_1, \ldots, x_n)$ are satisfied exactly by the same n-tuples $(a_1, \ldots, a_n)$ of elements that is

$$T \models \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n)).$$

Definition 1.2. We say that the theory T *defines* r *implicitly* if and only if it is not the case that there are two L'-models in which T holds, having the same elements and interpreting all symbols of L in the same way but interpreting the symbol r differently. This is also often expressed by

$$T, T' \models \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow r'(x_1, \ldots, x_n)),$$

where T' is exactly like T except that any occurrence of r is replaced by r', a predicate symbol of the same arity as r but not in L.

Notice that if a relation is explicitly definable then it is implicitly definable as well. What about the converse? The answer depends on the choice of the underlying logic. A logical system in which the converse holds is said to have the *Beth's definability property.*

In 1953, E.W. Beth [5] proved the following.

Theorem 1.3. *First order logic has the Beth's definability property.*

This thesis is on Beth's property. We will give a proof of Theorem 1.3 which is more detailed than in [5]. We will try to understand the proof by exhibiting a logic that does not have Beth's property and by looking at a proposed way to fix it. We will highlight the steps needed to achieve Beth property. In so doing, we hope to understand some of the crucial reasons why Beth property holds in first-order logic.

This thesis is organised as follows: Section 2 sets the notation and lists basic concepts and theorems from model theory of first order logic. Section 3 states and proves Beth definability theorem for first-order logic. Section 4 treats quantified modal logic in which Beth property does not hold. Section 5 looks at quantified hybrid logic, a logic devised to fix the failure of Beth's property in quantified modal logic.

Chapter 2

General notation and terminology

We assume that the reader is familiar with naive set theory and the basics of first-order logic. Throughout, we basically use the notation and terminology of [2] and [7]. To spare the reader looking into [2] and [7], we recall some basics.

2.1 Sets, relations and functions

Throughout, we "live" in Zermelo-Fraenkel set theory with the axiom of choice (ZFC). Right through, $\emptyset$ denotes the empty set. If x is a set, then $S(x)$ denotes its *successor* $x \cup \{x\}$. Recall that, according to von Neumann, a possible coding of the natural numbers in ZFC is:

$$0 = \emptyset, 1 = S(0), 2 = S(1), \ldots, n = S(n-1), \ldots.$$

It is left to the reader to check that this implies that

$$n = \{0, 1, 2, \ldots, n-1\} \tag{2.1.1}$$

for every natural number n. Thus $k \in n$ for every $k < n$ (where $<$ is the usual ordering of natural numbers). Throughout, ω denotes the set of all natural numbers (in von Neumann's sense).

If a and b are sets then the *ordered pair* with first member a and second b is denoted by $\langle a, b \rangle$. Recall that, in ZFC,

$$\langle a, b \rangle = \langle a_1, b_1 \rangle \text{ if and only if } a = a_1 \text{ and } b = b_1,$$

for every sets a, b, a_1, b_1.

Recall that a *binary relation* is defined to be a set of pairs. If R is a binary relation, then $\mathrm{Dom}(R)$ and $\mathrm{Rng}(R)$ denote its *domain* and *range*, respectively, that is,

$$\mathrm{Dom}(R) \stackrel{\text{def}}{=} \{x : \langle x, y \rangle \in R\} \text{ and } \mathrm{Rng}(R) \stackrel{\text{def}}{=} \{y : \langle x, y \rangle \in R\}.$$

If a binary relation f satisfies

$$\text{if } \langle x, y \rangle \in f \text{ and } \langle x, z \rangle \in f \text{ then } y = z$$

for every sets x, y, z then f is called a *function*. For any $x \in \mathrm{Dom}(f)$, $f(x)$ denotes the unique element y for which $\langle x, y \rangle \in f$. Instead of $f(x)$, we sometimes write fx or f_x. For a function f and sets A, B, "$f : A \longrightarrow B$" means that $\mathrm{Dom}(f) = A$ and $\mathrm{Rng}(f) \subseteq B$. If $f : A \longrightarrow B$ and $C \subseteq A$, then the restriction of f to C is the function $f \restriction C : C \longrightarrow B$ such that $f \restriction C(c) = f(c)$ for all $c \in C$. A function $f : A \longrightarrow B$ is called *surjective* or *onto* if $\mathrm{Rng}(f) = B$, *injective* or *one-to-one* if

$$f(a) = f(b) \implies a = b$$

for all $a, b \in A$; *bijective* if it is both surjective and injective.

Let A and B be sets. Then ${}^{A}B$ denotes the set of all functions from A into B, that is,

$$^{A}B \stackrel{\text{def}}{=} \{f : f \text{ is a function with } \operatorname{Dom}(f) = A \text{ and } \operatorname{Rng}(f) \subseteq B\}.$$

Thus ${}^{\emptyset}B = \{\emptyset\} = 1$ and ${}^{A}\emptyset = \emptyset = 0$ if $A \neq \emptyset$.

Sometimes we call functions *sequences*. In particular, we speak about finite sequences. If X is a set then f is called a *finite sequence* over X if $\operatorname{Dom}(f) \in \omega$ and $\operatorname{Rng}(f) \subseteq X$. According to (2.1.1) then

$$f : \{0, 1, 2, \ldots, n-1\} \longrightarrow X$$

for some $n \in \omega$. In this case, the finite sequence f can also be written as $\langle f_0, \ldots, f_{n-1}\rangle$. If $n = 0$, then $f = \emptyset$.

For any set X, X^* denotes the set of all *finite sequences over* X, defined as follows:

$$\begin{aligned} X^* &\stackrel{\text{def}}{=} \{f : \operatorname{Dom}(f) \in \omega \text{ and } \operatorname{Rng}(f) \subseteq X\} = \\ &= \bigcup \{{}^{n}X : n \in \omega\}. \end{aligned}$$

The elements of X^* are also called *words* over X, suggesting that, sometimes, the intuition behind a subset H of X^* is that H is a language over the alphabet X.

The concatenation $p^\frown q$ of two words $p := \langle a_1, \ldots, a_n\rangle$ and $q := \langle b_1, \ldots, b_k\rangle$ is just the two words written one after the other, that is,

$$p^\frown q = \langle a_1, \ldots, a_n\rangle^\frown\langle b_1, \ldots, b_k\rangle \stackrel{\text{def}}{=} \langle a_1, \ldots, a_n, b_1, \ldots, b_k\rangle.$$

We often simply write pq in place of $p^\frown q$, and we will use this notation extensively in the definitions below. We will often write just a in place of $\langle a\rangle$. Using these two conventions, we can write $a_1 \dots a_n$ in place of $\langle a_1, \dots, a_n \rangle$.

2.2 First-order logic

In this subsection we recall the definitions of formulas, models and satisfactions of first-order logic (FOL).

First we specify the alphabet from which we will build up our formulas. This alphabet will consist of the following parts:

- the so-called *logical symbols*: $\mathrm{LS} := \{\neg, \wedge, \exists, \doteq\}$,
- some auxiliary symbols (which could be omitted but their use makes our life easier): parentheses (and),
- some parameters: *non-logical symbols* (function symbols and relation symbols of Definition 2.1 below), and
- a set of variables.

Definition 2.1 (vocabulary)**.** We call a function t a *vocabulary* (or *signature* or *ranked alphabet* or *similarity type*) if conditions (i) and (ii) below hold.

(i) $\mathrm{Rng}(t) \subseteq \omega$,

(ii) $\mathrm{Dom}(t) = \mathrm{Fns}_t \sqcup \mathrm{Rls}_t$ for some sets Fns_t and Rls_t ($\sqcup$ denotes disjoint union).

The sets Fns_t and Rls_t are called the set of function symbols of t and the set of relation symbols of t, respectively. For any $s \in \mathrm{Dom}(t)$, $t(s)$ is called the *rank* or *arity* of s. If $s \in \mathrm{Fns}_t$ and $t(s) = 0$, then we call s a *constant symbol*.

The set $\mathrm{Dom}(t)$ is often called the set of non-logical symbols of an alphabet for FOL.

From now on, unless stated otherwise, t stands for an arbitrary but fixed vocabulary.

Let V be an arbitrary set satisfying $V \cap (\mathrm{Dom}(t) \cup \mathrm{LS}) = \emptyset$ (but arbitrary otherwise). We call V a set of *variables*.

Definition 2.2 (term and formula)**.** We define the set $\mathrm{Trm_t(V)}$ of *terms* of similarity type t with variables from V to be the smallest subset H of $(V \cup \mathrm{Fns_t})^*$ satisfying

(i) $V \subseteq H$ and

(ii) $\{f\tau_1 \dots \tau_n : f \in \mathrm{Fns}_t, t(f) = n \text{ and } \tau_1, \dots, \tau_n \in H\} \subseteq H$.

We define the set $\mathrm{Fml_t(V)}$ of *formulas* of similarity type t with variables from V to be the smallest subset H of $(V \cup \mathrm{Dom}(t) \cup \mathrm{LS})^*$ satisfying

(i) $\{r\tau_1 \dots \tau_n : r \in \mathrm{Rls}_t, t(r) = n, \text{ and } \tau_1, \dots, \tau_n \in \mathrm{Trm}_t(V)\} \cup$
$\cup \{\tau \doteq \sigma : \tau, \sigma \in \mathrm{Trm}_t(V)\} \subseteq H$ and

(ii) $\{\neg\varphi : \varphi \in H\} \cup \{\wedge\varphi\psi : \varphi, \psi \in H\} \cup$
$\cup \{\exists x\varphi : x \in V \text{ and } \varphi \in H\} \subseteq H$.

The formulas belonging to the left-hand-side of "$\subseteq$" in (i) are called *atomic formulas.*

Definition 2.3 (free and bound variables, sentence and theory)**.** Let $\varphi \in \mathrm{Fml}_t(V)$. We define the *free* and *bound variables* of φ inductively as follows:

- If φ is an atomic formula, the variable x is free in φ if and only if x occurs in φ. There is no bound variable in any atomic formula.

- If $\varphi = \neg\psi$, then x is free (respectively bound) in φ if and only if x is free (respectively bound)in ψ.

- If $\varphi = \psi \wedge \theta$, then x is free (respectively bound) in φ if and only if x is free (respectively bound) in either ψ or θ.

- If $\varphi = \exists y\psi$, then x is free in φ if and only if x is free in ψ and x and y are different symbols. Also, x is bound in φ if and only if x is y or x is bound in ψ.

If no variable occurs free in φ, then we say that φ is a *sentence* of $\mathrm{Fml}_t(V)$. A set of sentences T such that $T \subseteq \mathrm{Fml}_t(V)$ is called a *t-theory*.

The logical connectives $\neg, \wedge, \exists x$ are called, respectively, *negation, conjunction* and *existential quantifier*. For easier readability, we will often write $f(\tau_1, \ldots, \tau_n), r(\tau_1, \ldots, \tau_n)$ and $(\varphi \wedge \psi)$ in place of $f\tau_1 \ldots \tau_n, r\tau_1 \ldots \tau_n$ and $(\wedge\varphi\psi)$, respectively. If $\varphi \in \mathrm{Fml}_t(V)$, then we often refer to it as a *t-formula* or *t-sentence* if it is a sentence.

We will use the following standard abbreviations:

$$\begin{aligned}
(\varphi \vee \psi) &\text{ stands for } \neg(\neg\varphi \wedge \neg\psi),\\
(\varphi \rightarrow \psi) &\text{ stands for } \neg(\varphi \wedge \neg\psi),\\
(\varphi \leftrightarrow \psi) &\text{ stands for } (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi),\\
\forall v\varphi &\text{ stands for } \neg\exists\neg\varphi.
\end{aligned}$$

The derived logical connective $\vee$ is called *disjunction*, $\rightarrow$ is called *conditional* or *implication*, $\leftrightarrow$ is called *biconditional* or *equivalence*, and $\forall v$ is called *universal quantifier*.

Definition 2.4 (model and structure)**.** A *t-model* (or a *model of similarity type t* or a *t-structure*) $\mathfrak{M}$ is a pair $\langle \mathcal{U}(\mathfrak{M}), m\rangle$ satisfying the following conditions:

(i) $\mathcal{U}(\mathfrak{M})$ is a nonempty set called the *universe* of $\mathfrak{M}$,

(ii) m is a function such that

- $\mathrm{Dom}(m) = \mathrm{Dom}(t)$,
- if $f \in \mathrm{Fns}_t(V)$ and $t(f) = n$ then $m(f) : \mathcal{U}(\mathfrak{M})^n \longrightarrow \mathcal{U}(\mathfrak{M})$, and

- if $r \in \mathrm{Rls}_t(V)$ and $t(r) = n$ then $m(r) \subset \mathcal{U}(\mathfrak{M})^n$. For $n = 0$, we have $\mathcal{U}(\mathfrak{M})^n = \{\emptyset\}$.

For each symbol $s \in \mathrm{Dom}(t)$, we call $m(s)$ the interpretation of s in $\mathfrak{M}$ and we also denote by $s^{\mathfrak{M}}$.

Remark 2.5. In FOL, we make the notion of structure and model coincide. It will not be the case for Quantified Modal Logic. We make the distinction since for definability we reason in terms of models.

Definition 2.6 (valuation of variables, validity of formulas, and semantical consequence)**.** Let $\mathfrak{M}$ be a t-model and let V be an arbitrary set of variables for t. A function $k : V \longrightarrow \mathcal{U}(\mathfrak{M})$ is called a *valuation* of the variables from V in $\mathfrak{M}$.

Let k be an arbitrary but fixed valuation of the variables in $\mathfrak{M}$. We define when a t-formula is *true in* $\mathfrak{M}$ *at valuation* k *of the variables*, in symbols $\mathfrak{M} \vDash \varphi[k]$, by recursion, as follows. First we define the value $\tau^{\mathfrak{M}}[k]$ of any term $\tau \in \mathrm{Trm}_t(V)$ at k in $\mathfrak{M}$ as:

- $x^{\mathfrak{M}}[k] := k(x)$ if $x \in V$,
- $(f(\tau_1, \ldots, \tau_n))^{\mathfrak{M}}[k] :=$
  $f^{\mathfrak{M}}(\tau_1^{\mathfrak{M}}[k], \ldots, \tau_n^{\mathfrak{M}}[k])$ if $f \in \mathrm{Fns}_t, t(f) = n, \tau_1, \ldots, \tau_n \in \mathrm{Trm}_t(V)$.

Now

- for atomic formulas $r(\tau_1, \ldots, \tau_n)$,

$$\mathfrak{M} \vDash r(\tau_1, \ldots, \tau_n)[k] \overset{\text{def}}{\iff} \langle \tau_1^{\mathfrak{M}}[k], \ldots, \tau_n^{\mathfrak{M}}[k] \rangle \in r^{\mathfrak{M}},$$

  for atomic formulas $\tau \doteq \sigma$,

$$\mathfrak{M} \vDash (\tau \doteq \sigma)[k] \overset{\text{def}}{\iff} \tau^{\mathfrak{M}}[k] = \sigma^{\mathfrak{M}}[k],$$

- for negated formulas $\neg\varphi$,

$$\mathfrak{M} \models \neg\varphi[k] \stackrel{\text{def}}{\iff} \text{ it is not the case that } \mathfrak{M} \models \varphi[k](\text{ or } \mathfrak{M}\neg \models \varphi[k]),$$

- for conjunctions $(\varphi \wedge \psi)$,

$$\mathfrak{M} \models (\varphi \wedge \psi)[k] \stackrel{\text{def}}{\iff} \mathfrak{M} \models \varphi[k] \text{ and } \mathfrak{M} \models \psi[k],$$

- for quantified formulas $\exists x\varphi$,

$$\mathfrak{M} \models \exists x\varphi[k] \stackrel{\text{def}}{\iff} \mathfrak{M} \models \varphi[k'] \text{ for some valuation } k'$$
$$\text{such that } k \upharpoonright (V \setminus \{x\}) = k' \upharpoonright (V \setminus \{x\}).$$

By these, $\mathfrak{M} \models \varphi[k]$ has been defined for any t-formula φ.

We say that φ is *valid* in $\mathfrak{M}$ or $\mathfrak{M}$ is a *model* of φ, in symbols

$$\mathfrak{M} \models \varphi,$$

if $\mathfrak{M} \models \varphi[k]$ for every valuation $k : V \longrightarrow \mathcal{U}(\mathfrak{M})$. We say that φ is (logically) *valid*, in symbols $\models \varphi$, if $\mathfrak{M} \models \varphi$ for every t-model $\mathfrak{M}$.

We say that a t-model $\mathfrak{M}$ satisfies $\Sigma \subseteq \mathrm{Fml}_t(V)$, in symbols $\mathfrak{M} \models \Sigma$, if $\mathfrak{M} \models \varphi$, for all $\varphi \in \Sigma$.

If $\Sigma \subseteq \mathrm{Fml}_t(V)$ and $\varphi \in \mathrm{Fml}_t(V)$, then we say that φ is a *semantical consequence* of Σ, in symbols

$$\Sigma \models \varphi,$$

if for every t-model $\mathfrak{M}$, whenever $\mathfrak{M} \models \Sigma$ then $\mathfrak{M} \models \varphi$.

Notation 2.7. If $\mathfrak{M}$ is a t-model, $k : V \longrightarrow \mathcal{U}(\mathfrak{M})$, $\varphi \in \mathrm{Fml}_t(V)$, and $x_1, \ldots, x_n$ are all the variables occurring freely in φ and the order of these is fixed somehow, then instead of $\mathfrak{M} \vDash \varphi[k]$ we sometimes write $\mathfrak{M} \vDash \varphi[k(x_1), \ldots, k(x_n)]$ or $\mathfrak{M} \vDash \varphi(k(x_1), \ldots, k(x_n))$. If $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$, then $\mathfrak{M} \vDash \varphi(a_1, \ldots, a_n)$ is equivalent to $\mathfrak{M} \vDash \varphi(k(x_1), \ldots, k(x_n))$ for some valuation k such that $k(x_i) = a_i$ for $i = 1, \ldots, n$.

Convention 2.8. If $\varphi(x_1, \ldots, x_n) \in \mathrm{Fml}_t(V)$ and $\mathfrak{M}$ is a t-model, we notice that

$$\mathfrak{M} \vDash \varphi(x_1, \ldots, x_n) \text{ if and only if } \mathfrak{M} \vDash \forall x_1 \ldots \forall x_n \varphi(x_1, \ldots, x_n).$$

Therefore, we can always assume that whenever we write $\mathfrak{M} \vDash \varphi$, φ is a t-sentence. In the same manner, when we write $\mathfrak{M} \vDash T$ for $T \subseteq \mathrm{Fml}_t(V)$, we assume that T is a t-theory.

Next we list some basic concepts and theorems from model theory of FOL that we will refer to for our material on Beth definability property of FOL. Proofs of theorems are available in [7].

Definition 2.9 (expansion and reduct)**.** Let $t \subseteq t'$ be two vocabularies and let $\mathfrak{M}$ be a t-model and $\mathfrak{N}$ a t'-model.

We say that $\mathfrak{N}$ is an *expansion* of $\mathfrak{M}$ if $\mathcal{U}(\mathfrak{N}) = \mathcal{U}(\mathfrak{M})$ and for every symbol $s \in \mathrm{Dom}(t)$, $s^{\mathfrak{N}} = s^{\mathfrak{M}}$. If $\mathfrak{N}$ is an expansion of $\mathfrak{M}$ then we say that $\mathfrak{M}$ is a *reduct* of $\mathfrak{N}$ to the vocabulary t. Observe that every t'-model $\mathfrak{N}$ has exactly one reduct to t, which is denoted by $\mathfrak{N} \upharpoonright t$.

In the following definitions and properties, let t be a vocabulary and let $\mathfrak{M}$ and $\mathfrak{N}$ be t-models. We denote $Th(\mathfrak{M})$ the set of all t-sentences true in $\mathfrak{M}$. Any sentence, formula, and theory are to be understood as t-sentence, t-formula, and t-theory respectively.

Definition 2.10 (elementary equivalence)**.** We say that $\mathfrak{M}$ and $\mathfrak{N}$ are *elementary equivalent*, written $\mathfrak{M} \equiv \mathfrak{N}$, if $Th(\mathfrak{M}) = Th(\mathfrak{N})$.

Definition 2.11 (embedding and isomorphism)**.** An *embedding* from $\mathfrak{M}$ into $\mathfrak{N}$ is a function α from $\mathcal{U}(\mathfrak{M})$ into $\mathcal{U}(\mathfrak{N})$ such that:

(i) If c is a constant symbol in t then $\alpha(c^{\mathfrak{M}}) = c^{\mathfrak{N}}$.

(ii) If R is an n-ary relation symbol in t then for all $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$,

$$(a_1, \ldots, a_n) \in R^{\mathfrak{M}} \text{ if and only if } (\alpha(a_1), \ldots, \alpha(a_n)) \in R^{\mathfrak{N}}.$$

(iii) If f is an n-ary function symbol in t then for all $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$,

$$\alpha(f^{\mathfrak{M}}(a_1, \ldots, a_n)) = f^{\mathfrak{N}}(\alpha(a_1), \ldots, \alpha(a_n)).$$

An embedding is always injective. If in addition α is surjective, then we say that α is an *isomorphism* from $\mathfrak{M}$ onto $\mathfrak{N}$ and we write $\alpha : \mathfrak{M} \cong \mathfrak{N}$. We say that $\mathfrak{M}$ is isomorphic to $\mathfrak{N}$, and we write $\mathfrak{M} \cong \mathfrak{N}$, if there is a an isomorphism α from $\mathfrak{M}$ onto $\mathfrak{N}$.

Lemma 2.12. *If* $\mathfrak{M} \cong \mathfrak{N}$ *then* $\mathfrak{M} \equiv \mathfrak{N}$.

Definition 2.13 (submodel)**.** We say that $\mathfrak{M}$ is a *submodel* of $\mathfrak{N}$, and we write $\mathfrak{M} \subseteq \mathfrak{N}$ if the following conditions are satisfied:

(i) $\mathcal{U}(\mathfrak{M}) \subseteq \mathcal{U}(\mathfrak{N})$,

(ii) if c is a constant symbol in t then $c^{\mathfrak{N}} = c^{\mathfrak{M}}$,

(iii) if r is an n-ary relation symbol in t then $r^{\mathfrak{N}} \cap \mathcal{U}(\mathfrak{M})^n = r^{\mathfrak{M}}$ ($r^{\mathfrak{N}} = r^{\mathfrak{M}}$ if $n = 0$), and

(iv) if f is an n-ary function symbol in t then for all $a_i, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$, $f^{\mathfrak{N}}(a_1, \ldots, a_n) = f^{\mathfrak{M}}(a_1, \ldots, a_n)$.

Definition 2.14 (elementary submodel)**.** We say that $\mathfrak{M}$ is an *elementary submodel* of $\mathfrak{N}$, and we write $\mathfrak{M} \preccurlyeq \mathfrak{N}$, if $\mathfrak{M} \subseteq \mathfrak{N}$ and for every $n < \omega$, for every formula $\varphi(x_1, \ldots, x_n)$ and for all $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$ we have

$$\mathfrak{N} \vDash \varphi(a_1, \ldots, a_n) \text{ if and only if } \mathfrak{M} \vDash \varphi(a_1, \ldots, a_n)$$

Lemma 2.15. *If $\mathfrak{M} \preccurlyeq \mathfrak{N}$ then $\mathfrak{M} \equiv \mathfrak{N}$.*

Lemma 2.16 (Tarski-Vaught criterion)**.** *Let $\mathfrak{M} \subseteq \mathfrak{N}$. Then $\mathfrak{M} \preccurlyeq \mathfrak{N}$ if and only if for every $n < \omega$, for every formula $\varphi(x_1, \ldots, x_n, y)$ and for all $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M})$,*

$$\textit{if } \mathfrak{N} \vDash \exists y \varphi(a_1, \ldots, a_n, y) \textit{ then there is } b \in \mathcal{U}(\mathfrak{M}) \textit{ such that } \mathfrak{M} \vDash \varphi(a_1, \ldots, a_n, b).$$

Theorem 2.17 (completeness theorem)**.** *Let T be a theory and let φ be a sentence. Then,*

(i) T has a model if and only if T is consistent, and

(ii) $T \vDash \varphi$ if and only if $T \vdash \varphi$

Theorem 2.18 (compactness theorem)**.** *Let T be a theory and let φ be a sentence. Then,*

(i) T has a model if and only if every finite subset of T has a model, and

(ii) if $T \vDash \varphi$ then $U \vDash \varphi$ for some U finite subset of T.

Theorem 2.19 (deduction theorem)**.** *Let T be a theory, σ be a sentence and φ be a formula. Then,*

$$T \cup \{\sigma\} \vDash \varphi \textit{ if and only if } T \vDash \sigma \to \varphi.$$

Chapter 3

Beth property in First Order Logic

3.1 Beth definability theorem

We give a detailed proof of Beth's definability theorem using Craig's interpolation theorem. They both are results involving amalgamation of vocabularies. The proof is based on [7].

We will need the following definition.

Definition 3.1 (separability)**.** Let t_1 and t_2 be two similarity types such that $t_0 := t_1 \cap t_2$. Let θ be a t_0-sentence. Let T and U be a t_1-theory and a t_2-theory, respectively. We say that θ *separates* T and U if $T \vDash \theta$ and $U \vDash \neg\theta$. We say that T and U are *inseparable* if no t_0-sentence separates them.

Theorem 3.2 (Craig's interpolation theorem)**.** *Let φ be a t_1-sentence and ψ be a t_2-sentence. If $\varphi \vDash \psi$ then there exists a $t_1 \cap t_2$-sentence θ such that $\varphi \vDash \theta$ and $\theta \vDash \psi$.*

The sentence θ is called a Craig interpolant of φ and ψ.

Proof. Let t_1, t_2, φ, and ψ be as in the formulation of the theorem. Assume $\varphi \vDash \psi$ and that there is no interpolant of φ and ψ. We will derive a contradiction by showing that $\varphi \wedge \neg\psi$ has a model.

Let $t_0 = t_1 \cap t_2$. Let C be a countable infinite set of constant symbols not occurring in $t_1 \cup t_2$. Let $t'_i = t_i \cup C$, for $i = 0, 1, 2$.

Claim 3.2.1. *The t'_1-theory $\{\varphi\}$ and the t'_2-theory $\{\neg\psi\}$ are inseparable.*

Proof. For the sake of contradiction, assume that there exists a t'_0-sentence θ separating $\{\varphi\}$ and $\{\neg\psi\}$. Then we have $\varphi \vDash \theta$ and $\neg\psi \vDash \neg\theta$ or equivalently $\theta \vDash \psi$. We may assume that θ has the form $\theta'(c_1, \ldots, c_n)$, where $c_i \in C, i = 1, \ldots, n$ and $\theta'(x_1, \ldots, x_n)$ is a t_0-formula. Since φ and ψ do not contain any c_i for $i = 1, \ldots, n$, $\varphi \vDash \forall x_1 \ldots \forall x_n \theta'(x_1, \ldots, x_n)$ and $\forall x_1 \ldots \forall x_n \theta'(x_1, \ldots, x_n) \vDash \psi$, contradicting the fact that φ and ψ have no interpolant. $\square$

Let $\varphi_i, i < \omega$ and $\psi_i, i < \omega$ be enumerations of all t_1-sentences and all t_2-sentences respectively. We will construct two increasing sequences of theories

$$\{\varphi\} = T_0 \subseteq T_1 \subseteq T_2 \ldots$$
$$\{\neg\psi\} = U_0 \subseteq U_1 \subseteq U_2 \ldots$$

in the language of t'_1 and t'_2, respectively, such that conditions (1)–(3) below will be satisfied. For all $i < \omega$:

1. T_i and U_i are inseparable.
2. (a) if $T_i \cup \{\varphi_i\}$ and U_i are inseparable then $\varphi_i \in T_{i+1}$, and

   (b) if T_{i+1} and $U_i \cup \{\psi_i\}$ are inseparable then $\psi_i \in U_{i+1}$,
3. (a) if φ_i has the form $\exists x \sigma(x)$ and $\varphi_{i+1} \in T_i$ then $\sigma(c) \in T_{i+1}$ for some $c \in C$, and

   (b) if ψ_i has the form $\exists x \sigma(x)$ and $\psi_i \in U_{i+1}$then $\sigma(d) \in U_{i+1}$ for some $d \in C$.

Given T_i and U_i, T_{i+1} and U_{i+1} are constructed in the obvious way. We then have the following cases:

- $T_{i+1} = T_i$ (resp. $U_{i+1} = U_i$) if the condition in (2a) (resp. (2b)) is not satisfied,
- $T_{i+1} = T_i \cup \{\varphi_i, \sigma(c)\}$ (resp. $U_{i+1} = U_i \cup \{\psi_i, \sigma(d)\}$) if both the conditions in (2a) and (3a) (resp. (2b) and (3b)) are satisfied, and
- $T_{i+1} = T_i \cup \{\varphi_i\}$ (resp. $U_{i+1} = U_i \cup \{\psi_i\}$) if only the condition in (2a) (resp. (2b)) is satisfied.

For (3), c and d are chosen such that they did not occur in T_i, U_i, φ_i or ψ_i. In that way, inseparability is preserved.

Let $T_\omega = \bigcup_{i<\omega} T_i$ and $U_\omega = \bigcup_{i<\omega} U_i$. Since every T_i and U_i are finite theories for $i < \omega$, by the Compactness theorem, it follows that T_ω and U_ω are inseparable.

Claim 3.2.2. *The theories T_i and U_i are consistent for every $i \leqslant \omega$.*

Proof. Let $i \leqslant \omega$ be arbitrary but fixed. The theories T_i and U_i are inseparable. Therefore, they are both consistent. Since assume without loss of generality that T_i is not consistent. Then $T_i \vDash \neg\forall x(x = x)$ and $U_i \vDash \forall x(x = x)$, as $\forall x(x = x)$ is a tautology. But then we contradict the inseparability of T_i and U_i. □

Claim 3.2.3. *The theories T_ω and U_ω are maximal.*

Proof. Let σ be an arbitrary t_1'-sentence. We want to show that either $\sigma \in T_\omega$ or $\neg\sigma \in T_\omega$. Suppose for a contradiction that $\sigma \notin T_\omega$ and $\neg\sigma \notin T_\omega$. Then for some $i < \omega$, $\sigma = \varphi_i$ and $T_\omega \cup \{\varphi_i\}$ and U_ω are not inseparable. Hence, there exists a t_0'-sentence θ such that $T_\omega \cup \{\varphi_i\} \vDash \theta$ and $U_\omega \vDash \neg\theta$. By a similar argument, there is a t_0'-sentence θ' such that $T_\omega \cup \{\neg\varphi_i\} \vDash \theta'$ and $U_\omega \vDash \neg\theta'$. By the deduction theorem, we have

$$T_\omega \vDash \varphi_i \to \theta \text{ and } U_\omega \vDash \neg\theta$$

and

$$T_\omega \models \neg\varphi_i \to \theta' \text{ and } U_\omega \models \neg\theta'.$$

It follows that $T_\omega \models \theta \vee \theta'$ and $U_\omega \models \neg(\theta \vee \theta')$ contradicting the inseparability of T_ω and U_ω. Therefore, T_ω is maximal. In a similar way, one can show that U_ω is maximal. □

Claim 3.2.4. *The t_0'-theory $T_\omega \cap U_\omega$ is maximal consistent.*

Proof. Since $T_\omega \cap U_\omega \subseteq T_\omega$, it is consistent. We want to show that for every t_0'-sentence σ, either $\sigma \in T_\omega \cap U_\omega$ or $\neg\sigma \in T_\omega \cap U_\omega$. Let σ be a t_0'-sentence. By Claim 3.2.3, $\sigma \in T_\omega$ or $\neg\sigma \in T_\omega$ and $\sigma \in U_\omega$ or $\neg\sigma \in U_\omega$. Since T_ω and U_ω are inseparable, we cannot have $T_\omega \models \sigma$ and $U_\omega \models \neg\sigma$ or vice versa. Therefore, and by maximality of T_ω and U_ω, either $\sigma \in T_\omega \cap U_\omega$ or $\neg\sigma \in T_\omega \cap U_\omega$. □

Since T_ω is consistent, let $\mathfrak{N}_1$ be a t_1'-model such that $\mathfrak{N}_1 \models T_\omega$. Observe that for any constant symbol $e \in t_1$, for any n-ary function symbol $f \in t_1$ and any $c_1, \ldots, c_n \in C$, $\mathfrak{N}_1 \models \exists x(f(c_1, \ldots, c_n) = x)$ and $\mathfrak{N}_1 \models \exists x(e = x)$. Thus, by maximality of T_ω, $\exists x(f(c_1, \ldots, c_n) = x) \in T_\omega$ and $\exists x(e = x) \in T_\omega$. Using (3), we can then construct a submodel $\mathfrak{M}_1 \subseteq \mathfrak{N}_1$ such that

- $\mathcal{U}(\mathfrak{M}_1) = \{c^{\mathfrak{N}_1} : c \in C\}$,
- $e^{\mathfrak{M}_1} = e^{\mathfrak{N}_1}$ for every constant symbol in t_1',
- $R^{\mathfrak{M}_1} = R^{\mathfrak{N}_1} \cap (\mathcal{U}(\mathfrak{M}_1))^n$ for every n-ary relation symbol $R \in t_1$, and
- $f^{\mathfrak{M}_1}(a_1, \ldots, a_n) = f^{\mathfrak{N}_1}(a_1, \ldots, a_n)$ for every n-ary function symbol in t_1 and for all $a_1, \ldots, a_n \in \mathcal{U}(\mathfrak{M}_1)$.

Claim 3.2.5. *We have $\mathfrak{M}_1 \preccurlyeq \mathfrak{N}_1$, and in particular $\mathfrak{M}_1 \models T_\omega$.*

Proof. We use the Tarski-Vaught criterion (Lemma 2.16). Let $\mathfrak{N}_1 \vDash \exists y\varphi(c_1, \dots, c_n, y)$ for some t'_1-formula $\varphi(c_1, \dots, c_n, y)$, $c_1, \dots, c_n \in C$. By maximality of T_ω, $\exists y\varphi(c_1, \dots, c_n, y) \in T_\omega$ and by (3), $\varphi(c_1, \dots, c_n, c) \in T_\omega$ for some $c \in C$. $\square$

In the same way, let $\mathfrak{N}_2$ be a t'_2-model such that $\mathfrak{N}_2 \vDash U_\omega$. We make the following claim:

Claim 3.2.6. *There exists $\mathfrak{M}_2 \preccurlyeq \mathfrak{N}_2$ such that $\mathcal{U}(\mathfrak{M}_2) = \left\{c^{\mathfrak{N}_2} : c \in C\right\}$. In particular, $\mathfrak{M}_2 \vDash U_\omega$.*

Claim 3.2.7. *We have $\mathfrak{M}_1 \restriction t'_0 \cong \mathfrak{M}_2 \restriction t'_0$.*

Proof. We have $\mathfrak{M}_1 \restriction t'_0 \vDash T_\omega \cap U_\omega$ and $\mathfrak{M}_2 \restriction t'_0 \vDash T_\omega \cap U_\omega$. By maximality of $T_\omega \cap U_\omega$, for every t_0-formula $\varphi(x_1, \dots, x_n)$ and $c_1, \dots, c_n \in C$,

$$\mathfrak{M}_1 \restriction t'_0 \vDash \varphi(c_1, ..., c_n) \text{ if and only if } \mathfrak{M}_2 \restriction t'_0 \vDash \varphi(c_1, \dots, c_n) \ (*).$$

Let us denote $M_1 = \mathcal{U}(\mathfrak{M}_1 \restriction t'_0) = \mathcal{U}(\mathfrak{M}_1)$ and $M_2 = \mathcal{U}(\mathfrak{M}_2 \restriction t'_0) = \mathcal{U}(\mathfrak{M}_2)$. Let α be a function from M_1 into M_2 defined by $\alpha(c^{\mathfrak{M}_1}) = c^{\mathfrak{M}_2}$. By $(*)$, it is immediate to see that α is well defined and is an embedding of $\mathfrak{M}_1 \restriction t'_0$ into $\mathfrak{M}_2 \restriction t'_0$. Moreover, it is surjective hence an isomorphism. $\square$

Since $\mathfrak{M}_1 \restriction t'_0 \cong \mathfrak{M}_2 \restriction t'_0$ and $\mathfrak{M}_2 \vDash U_\omega$, we can expand $\mathfrak{M}_1 \restriction t'_0$ into a t'_2-model $\mathfrak{M}'_2$, such that $\mathfrak{M}'_2 \cong \mathfrak{M}_2$. We can then construct a $t'_1 \cup t'_2$-model $\mathfrak{M}$ such that it interprets t'_1 in the same manner as $\mathfrak{M}_1$ and t'_2 in the same manner as $\mathfrak{M}'_2$. Since $\varphi \in T_\omega$ and $\neg\psi \in U_\omega$, $\mathfrak{M} \vDash \varphi \wedge \neg\psi$, the contradiction we are looking for. $\square$

Definition 3.3 (implicit and explicit definition)**.** Let t be a vocabulary. Let $r, r' \notin t$ be two new n-ary relation symbols. Let $\Sigma(r)$ be a $t \cup \{r\}$-theory, and let $\Sigma(r')$ be the corresponding $t \cup \{r'\}$-theory formed by replacing r everywhere by r'.

We say that $\Sigma(r)$ *defines r implicitly* if

$$\Sigma(r) \cup \Sigma(r') \vDash \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow r'(x_1, \ldots, x_n)).$$

We say that $\Sigma(r)$ *defines r explicitly* if there exists a t-formula $\varphi(x_1, \ldots, x_n)$ such that

$$\Sigma(r) \vDash \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n)).$$

Theorem 3.4 (Beth definability theorem). *Let $\Sigma(r)$ be a $t \cup \{r\}$-theory for some vocabulary t and $r \notin t$. Then, $\Sigma(r)$ defines r explicitly if and only if it defines r implicitly.*

Proof. Suppose $\Sigma(r)$ defines r explicitly. Then by definition,

$$\Sigma(r) \vDash \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n)),$$

for a t-formula $\varphi(x_1, \ldots, x_n)$. But this is equivalent to

$$\Sigma(r') \vDash \forall x_1 \ldots \forall x_n (r'(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n)).$$

Combining the two, we have

$$\Sigma(r) \cup \Sigma(r') \vDash \forall x_1 \ldots \forall x_n (r(x_1, \ldots, x_n) \leftrightarrow \varphi(x_1, \ldots, x_n) \leftrightarrow r'(x_1, \ldots, x_n)).$$

Therefore, $\Sigma(r)$ defines r implicitly.

Conversely, suppose $\Sigma(r)$ defines r implicitly. Add new constants $c_1, \ldots, c_n$ to t. Then

$$\Sigma(r) \cup \Sigma(r') \vDash r(c_1, \ldots, c_n) \rightarrow r'(c_1, \ldots, c_n).$$

By the compactness theorem, there exists finite subsets $\Delta \subseteq \Sigma(r)$ and $\Delta' \subseteq \Sigma(r')$ such that

$$\Delta \cup \Delta' \vDash r(c_1, \ldots, c_n) \to r'(c_1, \ldots, c_n).$$

Let $\psi(r)$ be the conjunction of all $t \cup \{r\}$-sentences in Δ and $\psi(r')$ be the conjunction of all $t \cup \{r'\}$-sentences in Δ'. Then,

$$\psi(r) \wedge \psi(r') \vDash r(c_1, \ldots, c_n) \to r'(c_1, \ldots, c_n).$$

By the deduction theorem,

$$\psi(r) \wedge r(c_1, \ldots, c_n) \vDash \psi(r') \to r'(c_1, \ldots, c_n).$$

By Craig interpolation theorem, there exists a t-formula $\theta(x_1, \ldots, x_n)$ such that $\theta(c_1, \ldots, c_n)$ is a $t \cup \{c_1, \ldots, c_n\}$-sentence,

$$\psi(r) \wedge r(c_1, \ldots, c_n) \vDash \theta(c_1, \ldots, c_n), \tag{3.1.1}$$

and

$$\theta(c_1, \ldots, c_n) \vDash \psi(r') \to r'(c_1, \ldots, c_n). \tag{3.1.2}$$

By deduction theorem, (3.1.1) is equivalent to

$$\psi(r) \vDash r(c_1, \ldots, c_n) \to \theta(c_1, \ldots, c_n), \tag{3.1.3}$$

and (3.1.2) is equivalent to

$$\psi(r') \vDash \theta(c_1, \ldots, c_n) \to r'(c_1, \ldots, c_n)$$

which is again equivalent to

$$\psi(r) \vDash \theta(c_1, \dots, c_n) \to r(c_1, \dots, c_n). \tag{3.1.4}$$

Now (3.1.3) and (3.1.4) yield

$$\psi(r) \vDash r(c_1, \dots, c_n) \leftrightarrow \theta(c_1, \dots, c_n).$$

Because $c_1, \dots, c_n$ do not occur in $\psi(r)$,

$$\psi(r) \vDash \forall x_1 \dots \forall x_n (r(x_1, \dots, x_n) \leftrightarrow \theta(x_1, \dots, x_n)),$$

where $x_1, \dots, x_n$ are variables not occurring in $\theta(c_1, \dots, c_n)$. Since $\psi(r)$ is a conjunction of sentences in $\Sigma(r)$,

$$\Sigma(r) \vDash \forall x_1 \dots \forall x_n (r(x_1, \dots, x_n) \leftrightarrow \theta(x_1, \dots, x_n)).$$

As $\theta(x_1, \dots, x_n)$ is a t-formula, $\Sigma(r)$ defines r explicitly. We have proved Theorem 3.2. $\square$

Chapter 4

Beth property in Quantified Modal Logic

We will consider the quantified modal logic **S5** with constant domains denoted by **S5B** in [9].

4.1 Quantified modal logic

The language of quantified modal logic (QML) is obtained from the language of classical FOL by adding a unary operator $\Diamond$. Unless stated otherwise, we will basically follow the notation introduced in Section 2.

Logical symbols. The set LS of logical symbols is

$$\mathrm{LS} = \{\neg, \wedge, \exists, \doteq, \Diamond\}\,.$$

Vocabulary. We only consider vocabularies whose function symbols are of arity zero (constants). In the following, let t be an arbitrary but fixed vocabulary.

Formulas. The set $\mathrm{Fml}_t(V)$ of formulas is the smallest subset H of $(V \cup \mathrm{Dom}(t) \cup \mathrm{LS})^*$ satisfying

(i) $\{r\tau_1 \ldots \tau_n : r \in \mathrm{Rls}_t, t(r) = n,$ and $\tau_1, \ldots, \tau_n \in \mathrm{Trm}_t(V)\} \cup$
$\cup \{\tau \doteq \sigma : \tau, \sigma \in \mathrm{Trm}_t(V)\} \subseteq H$,

(ii) $\{\neg\varphi : \varphi \in H\} \cup \{\wedge\varphi\psi : \varphi, \psi \in H\} \cup$
$\cup \{\exists x\varphi : x \in V$ and $\varphi \in H\} \subseteq H$, and

(iii) $\{\Diamond\varphi : \varphi \in H\} \subseteq H$.

The modal operator $\Diamond$ is usually called *possibility.* For a formula φ, $\Box\varphi$ is the standard abbreviation of $\neg\Diamond\neg\varphi$. The modal operator $\Box$ is usually called *necessity.* The formula $\Diamond\varphi$ is usually read "*possibly* φ" and $\Box\varphi$ is usually read "*necessarily* φ".

Structure and Model. We define the following object over the vocabulary t: A *frame* is an ordered tuple $\langle W, R\rangle$ with W a nonempty set of states (or worlds) and R a binary relation on W. A *skeleton* is an ordered triple $\langle W, R, D\rangle$, with $\langle W, R\rangle$ a frame and D a function with domain W assigning to each state $w \in W$, a nonempty set D_w. Let $\bar{D}$ denote $\bigcup_{w\in W} D_w$.

A *structure* is an ordered quadruple $\mathfrak{S} \stackrel{\text{def}}{=} \langle W, R, D, m\rangle$ satisfying the following conditions:

(i) $\langle W, R, D\rangle$ is a skeleton,

(ii) m is a (interpretation) function such that

- $\mathrm{Dom}(m) = \mathrm{Dom}(t)$,
- if c is a constant, then $m(c) \in \bar{D}$, and

- if r is a relation symbol with $t(r) = n$, then $m(r) \subseteq W \times \bar{D} \times \cdots \times \bar{D}$ (n $\bar{D}$'s).

We use the standard notation $s^{\mathfrak{S}}$ to denote $m(s)$ for $s \in \mathrm{Dom}(t)$.

A *model* is a couple $\mathfrak{M} \stackrel{\mathrm{def}}{=} \langle \mathfrak{S}, w \rangle$ where $\mathfrak{S} \stackrel{\mathrm{def}}{=} \langle W, R, D, m \rangle$ is a structure and $w \in W$. Interpretation in $\mathfrak{M}$ is understood as interpretation in the underlying structure $\mathfrak{S}$, that is $s^{\mathfrak{M}} \stackrel{\mathrm{def}}{=} s^{\mathfrak{S}}$ for $s \in \mathrm{Dom}(t)$.

We say that a skeleton has *constant domains* if $D_w = D_v$ for all $w, v \in W$.

Truth. Let $\mathfrak{S} \stackrel{\mathrm{def}}{=} \langle W, R, D, m \rangle$ be a structure and $\mathfrak{M} \stackrel{\mathrm{def}}{=} \langle \mathfrak{S}, w \rangle$ be a model, where $w \in W$. Let V be an arbitrary set of variables. A *valuation* k is a function defined on V such that $k : V \longrightarrow \bar{D}$. For a valuation k and a term τ, we let $[k, m](\tau)$ denote $k(\tau)$ if τ is a variable, and $m(\tau)$ if it is a constant.

We define when a formula φ *is true in* $\mathfrak{M}$ *at valuation* k *of the variables*, in symbols $\mathfrak{M} \vDash \varphi[k]$, by recursion, with the following clauses:

- for atomic formulas $r(\tau_1, \ldots, \tau_n)$,

$$\mathfrak{M} \vDash r(\tau_1, \ldots, \tau_n)[k] \stackrel{\mathrm{def}}{\Longleftrightarrow} \langle w, [k, m](\tau_1), \ldots, [k, m](\tau_n) \rangle \in m(r),$$

  for atomic formulas $\tau_1 = \tau_2$,

$$\mathfrak{M} \vDash (\tau_1 = \tau_2)[k] \stackrel{\mathrm{def}}{\Longleftrightarrow} [k, m](\tau_1) = [k, m](\tau_2),$$

- for negated formulas $\neg\varphi$,

$$\mathfrak{M} \vDash \neg\varphi[k] \stackrel{\mathrm{def}}{\Longleftrightarrow} \mathfrak{M} \nvDash \varphi[k],$$

- for conjunction $\varphi \wedge \psi$,

$$\mathfrak{M} \vDash (\varphi \wedge \psi)[k] \stackrel{\text{def}}{\iff} \mathfrak{M} \vDash \varphi[k] \text{ and } \mathfrak{M} \vDash \psi[k],$$

- for modal formulas $\Diamond\varphi$,

$$\mathfrak{M} \vDash \Diamond\varphi[k] \stackrel{\text{def}}{\iff} \text{ there exists a } v \in W$$
$$\text{such that } wRv \text{ and } \langle \mathfrak{S}, v \rangle \vDash \varphi[k], \text{ and}$$

- for quantified formulas $\exists x\varphi$,

$$\mathfrak{M} \vDash \exists x\varphi[k] \stackrel{\text{def}}{\iff} \mathfrak{M} \vDash \varphi[k'] \text{ for some valuation } k'$$
$$\text{such that } k \upharpoonright (V \setminus \{x\}) = k' \upharpoonright (V \setminus \{x\}) \text{ and } \operatorname{Rng}(k') \subseteq D_w.$$

Let F be a class of skeleton, and φ a formula. We say that φ is F-valid, in symbols $\vDash_{\mathsf{F}} \varphi$ if for every structure $\mathfrak{S}$ on every skeleton from F, $\langle \mathfrak{S}, w \rangle \vDash \varphi[k]$ holds for every w and k. Validity of φ on the class of all skeletons is denoted by $\vDash \varphi$.

Semantical consequence. We say that a model $\mathfrak{M}$ satisfies a theory Σ, in symbols $\mathfrak{M} \vDash \Sigma$, if $\mathfrak{M} \vDash \varphi$, for all $\varphi \in \Sigma$.

We say that φ is a *semantical consequence* of Σ, in symbols

$$\Sigma \vDash \varphi,$$

if for every t-model $\mathfrak{M}$, whenever $\mathfrak{M} \vDash \Sigma$ then $\mathfrak{M} \vDash \varphi$. Here we use the notion of model not structure. This semantical consequence is often called local semantical consequence in the literature.

4.2 Failure of Beth property for quantified S5B

An **S5**-structure $\mathfrak{S} = \langle W, R, D, m\rangle$ is one in which R is $W \times W$. We denote a quantified modal logic **S5** with constant domains by **S5B**. We will give the counterexample to Beth's definability theorem for quantified **S5B** constructed in [9].

Let $\mathfrak{S} = \langle W, R, D, m\rangle$ be an **S5B**-structure for a vocabulary t without constant symbols and let $w \in W$.

Define $\mathfrak{S}_w$ to be the first-order structure $\langle \bar{D}, m_w\rangle$ such that

$$m_w(r) = \left\{\langle a_1, \ldots, a_n\rangle \in \bar{D}^n : \langle w, a_1, \ldots, a_n\rangle \in m(r)\right\}$$

for any n-ary relation symbol r in t.

Definition 4.1 (Isomorphism in **S5B**)**.** Let $\mathfrak{S} = \langle W, R, D, m\rangle$ and $\mathfrak{T} = \langle V, S, E, n\rangle$ be two **S5B**-structures. We say that σ is an *isomorphism from* $\mathfrak{S}$ *onto* $\mathfrak{T}$, in symbols $\sigma : \mathfrak{S} \cong \mathfrak{T}$, if σ is a bijection from $\bar{D}$ onto $\bar{E}$ such that

(i) For every $w \in W$, there exists $v \in V$ such that $\sigma : \mathfrak{S}_w \cong \mathfrak{T}_v$, and

(ii) for every $v \in V$, there exists $w \in W$ such that $\sigma : \mathfrak{S}_w \cong \mathfrak{T}_v$.

Lemma 4.2. *Let $\mathfrak{S} = \langle W, R, D, m\rangle$ and $\mathfrak{T} = \langle V, S, E, n\rangle$ be two **S5B**-structures. Suppose that $\sigma : \mathfrak{S} \cong \mathfrak{T}$ and for $w \in W$, $v \in V$, $\sigma : \mathfrak{S}_w \cong \mathfrak{T}_v$. Then, for any formula $\varphi(x_1, \ldots, x_n)$ with free variables $x_1, \ldots, x_n$ and for any tuple $\langle a_1, \ldots, a_n\rangle \in \bar{D}^n$,*

$$\langle \mathfrak{S}, w\rangle \vDash \varphi(a_1, \ldots, a_n) \text{ if and only if } \langle \mathfrak{T}, v\rangle \vDash \varphi(\sigma(a_1), \ldots, \sigma(a_n)).$$

Proof. We will prove by induction on the complexity of φ. The cases $\varphi(x_1, x_2) = (x_1 \doteq x_2)$, $\varphi = \psi \wedge \chi$, and $\varphi = \neg\psi$ are trivial. The remaining cases are as follows:

- $\varphi(x_1, \ldots, x_n) = r(x_1, \ldots, x_n)$ for r n-ary relation symbol. We have

$$\begin{aligned}
\langle \mathfrak{S}, w\rangle \vDash r(a_1, \ldots, a_n) &\iff \langle w, a_1, \ldots, a_n\rangle \in r^{\mathfrak{S}} && \text{[by definition]}\\
&\iff \langle a_1, \ldots, a_n\rangle \in r^{\mathfrak{S}_w} && \text{[by definition of } \mathfrak{S}_w]\\
&\iff \langle \sigma(a_1), \ldots, \sigma(a_n)\rangle \in r^{\mathfrak{T}_v} && [\sigma : \mathfrak{S}_w \cong \mathfrak{T}_v]\\
&\iff \langle v, \sigma(a_1), \ldots, \sigma(a_n)\rangle \in r^{\mathfrak{T}} && \text{[by definition of } \mathfrak{T}_v]\\
&\iff \langle \mathfrak{T}, v\rangle \vDash r(\sigma(a_1), \ldots, \sigma(a_n)). && \text{[by definition]}
\end{aligned}$$

- $\varphi(x_1, \ldots, x_n) = \Diamond\psi(x_1, \ldots, x_n)$. Assume that $\langle \mathfrak{S}, w\rangle \vDash \Diamond\psi(x_1, \ldots, x_n)$. Then there exists $w' \in W$ such that $\langle \mathfrak{S}, w'\rangle \vDash \psi(x_1, \ldots, x_n)$. Since $\sigma : \mathfrak{S} \cong \mathfrak{T}$, there exists $v' \in V$ such that $\sigma : \mathfrak{S}_{w'} \cong \mathfrak{T}_{v'}$. By induction hypothesis, $\langle \mathfrak{T}, v'\rangle \vDash \psi(\sigma(a_1), \ldots, \sigma(a_n))$. Therefore $\langle \mathfrak{T}, v\rangle \vDash \Diamond\psi(\sigma(a_1), \ldots, \sigma(a_n))$. Conversely, since $\sigma^{-1} : \mathfrak{T} \cong \mathfrak{S}$, $\langle \mathfrak{T}, v\rangle \vDash \Diamond\psi(\sigma(a_1), \ldots, \sigma(a_n))$ implies $\langle \mathfrak{S}, w\rangle \vDash \Diamond\psi(a_1, \ldots, a_n)$.

- $\varphi(x_1, \ldots, x_n) = \exists x\psi(x, x_1, \ldots, x_n)$. We have

$$\begin{aligned}
\langle \mathfrak{S}, w\rangle \vDash \exists x\psi(x, x_1, \ldots, x_n) &\iff \langle \mathfrak{S}, w\rangle \vDash \psi(a, a_1, \ldots, a_n) \text{ for some } a \in \bar{D}\\
&\iff \langle \mathfrak{T}, v\rangle \vDash \psi(\sigma(a), \sigma(a_1), \ldots, \sigma(a_n))\\
&\qquad \text{[by induction hypothesis]}\\
&\iff \langle \mathfrak{T}, v\rangle \vDash \exists x\psi(x, \sigma(a_1), \ldots, \sigma(a_n)).\\
&\qquad \text{[semantics of } \exists x]
\end{aligned}$$

□

Lemma 4.3. *Let $\mathfrak{S} = \langle W, R, D, m\rangle$ and $\mathfrak{T} = \langle V, S, E, n\rangle$ be two **S5B**-structures. Let $w \in W$ and $v \in V$. Suppose that $\rho : \mathfrak{S}_w \cong \mathfrak{T}_v$. Suppose also that for every finite ρ' such that $\rho' \subseteq \rho$, there exists σ containing ρ' such that $\sigma : \mathfrak{S} \cong \mathfrak{T}$. Then, for any formula*

$\varphi(x_1, \ldots, x_n)$ *with free variables* $x_1, \ldots, x_n$ *and for any tuple* $\langle a_1, \ldots, a_n \rangle \in \bar{D}^n$,

$$\langle \mathfrak{S}, w \rangle \vDash \varphi(a_1, \ldots, a_n) \text{ if and only if } \langle \mathfrak{T}, v \rangle \vDash \varphi(\rho(a_1), \ldots, \rho(a_n)).$$

Proof. We will prove by induction on the complexity of φ. Again, the cases $\varphi(x_1, x_2) = (x_1 \doteq x_2)$, $\varphi = (\psi \wedge \chi)$, and $\varphi = \neg\psi$ are trivial. For the other cases, we have

- $\varphi(x_1, \ldots, x_n) = r(x_1, \ldots, x_n)$. Same proof as in Lemma 4.2.
- $\varphi(x_1, \ldots, x_n) = \Diamond\psi(x_1, \ldots, x_n)$. Assume that $\langle \mathfrak{S}, w \rangle \vDash \Diamond\psi(a_1, \ldots, a_n)$. Then there exists $w' \in W$ such that $\langle \mathfrak{S}, w' \rangle \vDash \psi(a_1, \ldots, a_n)$. By the assumption, there is σ such that $\sigma : \mathfrak{S} \cong \mathfrak{T}$ and $\sigma \restriction \{a_1, \ldots, a_n\} = \rho'$. Therefore, there exists $v' \in V$ such that $\sigma : \mathfrak{S}_{w'} \cong \mathfrak{T}_{v'}$. By Lemma 4.2, $\langle \mathfrak{T}, v' \rangle \vDash \psi(\sigma(a_1), \ldots, \sigma(a_n))$. Since ρ and σ agree on $a_1, \ldots, a_n$, we have $\langle \mathfrak{T}, v' \rangle \vDash \psi(\rho(a_1), \ldots, \rho(a_n))$. Hence, $\langle \mathfrak{T}, v \rangle \vDash \Diamond\psi(\rho(a_1), \ldots, \rho(a_n))$. Conversely, since $\rho^{-1} : \mathfrak{T} \cong \mathfrak{S}$, $\langle \mathfrak{T}, v \rangle \vDash \Diamond\psi(\rho(a_1), \ldots, \rho(a_n))$ implies $\langle \mathfrak{S}, w \rangle \vDash \Diamond\psi(a_1, \ldots, a_n)$.
- $\varphi(x_1, \ldots, x_n) = \exists x \psi(x, x_1, \ldots, x_n)$. Same proof as in Lemma 4.2.

□

Theorem 4.4. *The quantified* ***S5B*** *does not have the Beth property.*

Proof. We give the counterexample in [9]. Let $T = \{\varphi_1, \varphi_2\}$ be the theory such that

$$\varphi_1 \stackrel{\text{def}}{=} p \to \Diamond\forall x(rx \to \Box(p \to \neg rx))$$

and

$$\varphi_2 \stackrel{\text{def}}{=} \neg p \to \Box\exists x(rx \wedge \Box(\neg p \to rx)).$$

Here p is a proposition (nullary relation symbol) and r is an unary relation symbol.

Claim 4.4.1. *The proposition p is implicitly definable in T.*

Proof. Let $\mathfrak{S} = \langle W, R, D, m\rangle$ be a **S5B**-structure and let $\langle \mathfrak{S}, w_0\rangle \vDash T$ for some $w_0 \in W$, that is $\langle \mathfrak{S}, w_0\rangle \vDash \varphi_i$ for $i = 1, 2$. Given any world w of W, let $\bar{r}_w$ be

$$\bar{r}_w \stackrel{\text{def}}{=} \left\{a \in \bar{D} : \langle w, a\rangle \in m(r)\right\}.$$

We have $a \in \bar{r}_w$ if and only if $\langle \mathfrak{S}, w\rangle \vDash ra$.

Since $\langle \mathfrak{S}, w_0\rangle \vDash \varphi_1$, $\langle \mathfrak{S}, w_0\rangle \vDash p$ implies that for some $w \in W$, $\bar{r}_w$ is disjoint from $\bar{r}_{w_0}$. In fact, we have $\langle \mathfrak{S}, w_0\rangle \vDash \Diamond\forall x(rx \to \Box(p \to \neg rx))$ which means that there exists $w \in W$ such that for all $a \in \bar{D}$, if $\langle \mathfrak{S}, w\rangle \vDash ra$ then $\langle \mathfrak{S}, w\rangle \vDash \Box(p \to \neg ra)$, in particular $\langle \mathfrak{S}, w_0\rangle \vDash p \to \neg ra$. Thus, for all $a \in \bar{D}$, if $a \in \bar{r}_w$, then $a \notin \bar{r}_{w_0}$, that is $\bar{r}_w \cap \bar{r}_{w_0} = \emptyset$.

Since $\langle \mathfrak{S}, w_0\rangle \vDash \varphi_2$, $\langle \mathfrak{S}, w_0\rangle \nvDash p$ implies that for no w is $\bar{r}_w$ disjoint from $\bar{r}_{w_0}$. In fact, we have $\langle \mathfrak{S}, w_0\rangle \vDash \Box\exists x(rx \wedge \Box(\neg p \to rx))$. That means for all $w \in W$, there exists $a \in \bar{D}$ such that $a \in \bar{r}_w$ and $\langle \mathfrak{S}, w\rangle \vDash \Box(\neg p \to ra)$, in particular $\langle \mathfrak{S}, w_0\rangle \vDash \neg p \to ra$. Thus, for all $w \in W$, $a \in \bar{r}_w \cap \bar{r}_{w_0}$.

Hence, $\langle \mathfrak{S}, w_0\rangle \vDash p$ if and only if

$$\text{there exists } w \in W \text{ such that } \bar{r}_w \cap \bar{r}_{w_0} = \emptyset, \tag{4.2.1}$$

and the implicit definability of p follows. □

Claim 4.4.2. *The proposition p is not explicitly definable in T.*

Proof. We construct an **S5B**-structure $\mathfrak{S} = \langle W, R, D, m\rangle$ of vocabulary $\{r\}$. Let $\bar{D}$ be the set of natural numbers $\mathbb{N}$. A permutation π on $\bar{D}$ is called *essentially finite* if $\left\{a \in \bar{D} : \pi(a) \neq a\right\}$ is finite. Let W be the set

$$\left\{\langle i, \pi\rangle : i = 0, 1, 2 \text{ and } \pi \text{ is an essentially finite permutation on } \bar{D}\right\}.$$

Let $\mathbb{O}$ be set of odd natural numbers and let $\mathbb{E}$ be the set of even natural numbers. Define $\bar{r}_{\langle 0,\pi\rangle} \stackrel{\text{def}}{=} \pi(\mathbb{N})$, $\bar{r}_{\langle 1,\pi\rangle} \stackrel{\text{def}}{=} \pi(\mathbb{O})$, and $\bar{r}_{\langle 2,\pi\rangle} \stackrel{\text{def}}{=} \pi(\mathbb{E})$.

Let ι be the identity permutation on $\bar{D}$, and write $w_i = \langle i, \iota\rangle$, for $i = 0, 1, 2$. Let ρ be any permutation on $\bar{D}$ such that $\rho(\mathbb{N}) = \mathbb{O}$. Then, $\rho : \mathfrak{S}_{w_0} \cong \mathfrak{S}_{w_1}$. In fact, for an arbitrary $a \in \mathbb{N}$,

$$\begin{aligned}
\mathfrak{S}_{w_0} \vDash ra &\iff \langle \mathfrak{S}, w_0\rangle \vDash ra \\
&\iff a \in \mathbb{N} \\
&\iff \rho(a) \in \mathbb{O} \\
&\iff \langle \mathfrak{S}, w_1\rangle \vDash r\rho(a) \\
&\iff \mathfrak{S}_{w_1} \vDash r\rho(a).
\end{aligned}$$

Also for every finite ρ' such that $\rho' \subseteq \rho$, there exists σ containing ρ' such that $\sigma : \mathfrak{S} \cong \mathfrak{S}$. For take any finite $\rho' \subseteq \rho$. Clearly, there is an essentially finite permutation σ such that $\rho' \subseteq \sigma$. Now, we have for $i = 0, 1, 2$, $\sigma : \mathfrak{S}_{\langle i,\pi\rangle} \cong \mathfrak{S}_{\langle i,\sigma\circ\pi\rangle}$ since $a \in \pi(X)$ if and only if $\sigma(a) \in \sigma(\pi(X))$ with $X = \mathbb{N}, \mathbb{O}, \mathbb{E}$. Therefore $\sigma : \mathfrak{S} \cong \mathfrak{S}$. By Lemmma 4.3, we have

$$\langle \mathfrak{S}, w_0\rangle \vDash \theta \text{ if and only if } \langle \mathfrak{S}, w_1\rangle \vDash \theta \text{ for any } \{r\}\text{-sentence } \theta. \tag{4.2.2}$$

Let $\mathfrak{T}_0$ and $\mathfrak{T}_1$ be two expansions of $\mathfrak{S}$ to the vocabulary $\{r, p\}$ such that $p^{\mathfrak{T}_0} = W \setminus \{w_0\}$ and $p^{\mathfrak{T}_1} = \{w_1\}$.

We have $\langle \mathfrak{T}_0, w_0\rangle \vDash T$. Firstly, since $\langle \mathfrak{T}_0, w_0\rangle \nvDash p$, $\langle \mathfrak{T}_0, w_0\rangle \vDash \varphi_1$. Secondly, we show that $\langle \mathfrak{T}_0, w_0\rangle \vDash \varphi_2$. Since $\langle \mathfrak{T}_0, w_0\rangle \vDash \neg p$, we need to show that $\langle \mathfrak{T}_0, w_0\rangle \vDash \Box\exists x(rx \wedge \Box(\neg p \to rx))$. That is, for each $w \in W$, we must find an $a \in \bar{D}$ such that (1) $a \in \bar{r}_w$ and (2) whenever $w' \notin p^{\mathfrak{T}_0}$ then $a \in \bar{r}_{w'}$. The condition of (2) gives $w' = w_0$. Since $\bar{r}_{w_0} = \bar{D}$, it is always possible to find such an a for each w. Hence, $\langle \mathfrak{T}_0, w_0\rangle \vDash \varphi_2$.

We also have $\langle \mathfrak{T}_1, w_1\rangle \vDash T$. Firstly, since $\langle \mathfrak{T}_1, w_1\rangle \nvDash \neg p$, $\langle \mathfrak{T}_1, w_1\rangle \vDash \varphi_2$. Secondly, we want to show that $\langle \mathfrak{T}_1, w_1\rangle \vDash \varphi_1$. Since $\langle \mathfrak{T}_1, w_1\rangle \vDash p$, we need to show that $\langle \mathfrak{T}_1, w_1\rangle \vDash \Diamond\forall x(rx \to \Box(p \to \neg rx))$. In other words, we have to find a world $w' \in W$ such that for all $a \in \bar{D}$, if $a \in \bar{r}_{w'}$, then $a \notin \bar{r}_{w_1}$. It suffices to choose w' to be w_2, for $\bar{r}_{w_1} = \mathbb{O}$, $\bar{r}_{w_2} = \mathbb{E}$ and $\mathbb{O} \cap \mathbb{E} = \emptyset$.

For the sake of contradiction, suppose that $T \vDash p \leftrightarrow \theta$ for a $\{r\}$-sentence θ. Since $\langle \mathfrak{T}_1, w_1\rangle \vDash T$ and $\langle \mathfrak{T}_1, w_1\rangle \vDash p$, $\langle \mathfrak{T}_1, w_1\rangle \vDash \theta$. Therefore, $\langle \mathfrak{S}, w_1\rangle \vDash \theta$ and by (4.2.2), $\langle \mathfrak{S}, w_0\rangle \vDash \theta$. But then, $\langle \mathfrak{T}_0, w_0\rangle \vDash \theta$, and since $\langle \mathfrak{T}_0, w_0\rangle \vDash T$, $\langle \mathfrak{T}_0, w_0\rangle \vDash p$, a contradiction.

□

Therefore, the quantified **S5B** does not have the Beth property. □

There is a standard translation of modal language into classical language. We give a translation inspired from [8] for quantified modal logic.

Let t be a vocabulary. Let t^* be a vocabulary such that it contains

- each constant of t plus a new constant w_0^*,
- for each n-ary relation symbol r of t an $(n+1)$-ary relation symbol r^*,
- two new unary relation symbol D^* and W^*, and
- two new binary relation symbol R^* and E^*.

Reserve one variable w^* of t^* and enumerate the others. If τ is the n-th variable of t, let τ^* be the n-th variable of t^* in the enumeration; and if c is a constant of t let c^* be c. Each formula φ of t may then be translated into a formula φ^* of t^* by means of the following clauses:

(i) (a) $(r\tau_1 \dots \tau_n)^* := W^*w^* \wedge D^*\tau_1^* \wedge \dots \wedge D^*\tau_n^* \wedge r^*w^*\tau_1^* \dots \tau_n^*$

(b) $(\tau_1 = \tau_2)^* := (\tau_1^* = \tau_2^*)$

(ii) $(\neg\varphi)^* := \neg(\varphi)^*$

(iii) $(\varphi \wedge \psi)^* := (\varphi^* \wedge \psi^*)$

(iv) $(\exists x\varphi)^* := W^*w^* \wedge \exists x^*(D^*x^* \wedge E^*w^*x^* \wedge \varphi^*)$

(v) $(\Diamond\varphi)^* := \exists w^*(W^*w^* \wedge R^*w_0^*w^* \wedge \varphi^*)$.

Let $w_0 \in W$ and $\mathfrak{M} := \langle \mathfrak{S}, w_0 \rangle$ be a **S5B**-model.

With each **S5B**-model $\mathfrak{M} := \langle \mathfrak{S}, w_0 \rangle$, where $\mathfrak{S} := \langle W, R, D, m \rangle$ is a **S5B**-structure of vocabulary t and $w_0 \in W$, associate a first order model $\mathfrak{M}^* := \langle W \cup \bar{D}, m^* \rangle$ of similarity type t^*. The interpretation m^* is defined as follows:

- if c is a constant in t, then $m^*(c^*) := m(c)$,
- $m^*(w_0^*) := w_0$,
- if r is an n-ary relation symbol in t, then $m^*(r^*) := m(r)$,
- $m^*(D^*) := \bar{D}$,
- $m^*(W^*) := W$,
- $m^*(R^*) := R$, and
- $m^*(E^*) := \big\{ \langle w, a \rangle \in W \times \bar{D} : a \in D_w \big\}$.

By straightforward induction and by the very definition of truth in model for QML, we have for any t-sentence φ

$$\mathfrak{M} \vDash \varphi \iff \mathfrak{M}^* \vDash \varphi^*[w^* \leftarrow w_0^*],$$

where the notation $[w^* \leftarrow w_0^*]$ means "replace every free occurrence of w^* with w_0^*".

In the counterexample given in the proof of Theorem 4.4, for a model $\langle \mathfrak{S}, w_0 \rangle$, $\langle \mathfrak{S}, w_0 \rangle \vDash p$ if and only if (4.2.1) is satisfied. But since p is not explicitly definable, there is no modal formula to express (4.2.1). In the next section, we will consider an extension of QML, expressive enough such that Beth property holds.

Chapter 5

Beth property in Quantified Hybrid Logic

Hybrid logics are extension of modal logics in which it is possible to reason about what happens at particular worlds. In modal logic, one cannot name worlds nor quantify over them. Starting with the vocabularies of QML, hybrid logic uses four tools: nominals, satisfaction operators, the $\downarrow$-binder to name worlds and to assert that a formula is true at a named world, and variables over worlds. Let the language of quantified hybrid logic (QHL) be the expansion of QML with these four tools. This section is based on [3]. Hybrid structures are expansions of modal structures.

5.1 Quantified Hybrid Logic

Nominals and satisfaction operators. Let NOM be a set of nullary relation symbols or propositional symbols distinct from any propositional symbols already in the vocabulary. These new symbols are called *nominals*. They can be compared with constants. While constants are name for individuals in universes, nominals are name for worlds. However, we notice that unlike constants, nominals are formulas. We also introduce the new satisfaction operators $@_n$ indexed by nominals. We then have two new types of formulas:

- for $n \in \mathsf{NOM}$, n is a formula and
- if φ is a formula and $n \in \mathsf{NOM}$, then $@_n\varphi$ is also a formula.

The formula $@_n\varphi$ is read "at n, φ" and intuitively it means that formula φ holds at the world named n.

Let $\mathfrak{S} := \langle W, R, D, m\rangle$ be a quantified modal structure. For $n \in \mathsf{NOM}$, $m(n) \subseteq W$. We impose for a nominal n to be interpreted as a singleton, that is for every $n \in \mathsf{NOM}$, there exists a unique $w \in W$ such that $m(n) = \{w\}$. Following the terminology of [3], the unique state w is called the denotation of n in $\mathfrak{S}$.

For satisfaction in models for formula involving the satisfaction operators $@_n$, we add the following clause:

- $\langle \mathfrak{S}, w\rangle \vDash @_n\varphi[k] \overset{\text{def}}{\iff} \langle \mathfrak{S}, \bar{n}\rangle \vDash \varphi[k]$,

where $\bar{n}$ is the denotation of n in $\mathfrak{S}$, $w \in W$ and k a valuation.

The ↓-binder. Let WVAR be a set of variables disjoint from the variables we already have. Those new variables will range over worlds. Again, unlike the already existing variables, those new variables are formulas. The ↓-binder is the analogous of $\exists$. We then have the following new types of formulas:

- every $\alpha \in \mathsf{WVAR}$ is a formula,
- if φ is a formula and $\alpha \in \mathsf{WVAR}$, then $@_\alpha\varphi$ is a formula, and
- if φ is a formula and $\alpha \in \mathsf{WVAR}$, then $\downarrow \alpha.\, \varphi$ is a formula.

In order to define truth in a model for formulas involving the newly introduced symbols, we extend valuation to elements of WVAR. Therefore, if $\alpha \in \mathsf{SVAR}$, and if k is a valuation then $k(\alpha) \in W$. Now, let $\mathfrak{S} := \langle W, R, D, m \rangle$ be a structure, $w \in W$, $\alpha \in \mathsf{WVAR}$, and k a valuation. We have the following clauses:

- $\langle \mathfrak{S}, w \rangle \models \alpha[k] \overset{\text{def}}{\iff} k(a) = w$,
- $\langle \mathfrak{S}, w \rangle \models @_\alpha \varphi[k] \overset{\text{def}}{\iff} \langle \mathfrak{S}, k(\alpha) \rangle \models \varphi[k]$, and
- $\langle \mathfrak{S}, w \rangle \models \downarrow \alpha.\, \varphi[k] \overset{\text{def}}{\iff} \langle \mathfrak{S}, w \rangle \models \varphi[k^\alpha_w]$,

where k^α_w is the assignment which differs from k only in that $k^\alpha_w(\alpha) = w$.

We give the additional clauses needed for a standard translation of any formula in QHL. For that we first need to expand the vocabulary t^* with new constants: for each nominal n add a constant $\tilde{n}$ in t^*. For variables, if $\alpha \in WVAR$, then add $\tilde{\alpha}$ as variable in t^*. Now, for $\alpha \in \mathsf{WVAR}$ and $n \in \mathsf{NOM}$,

- $(\downarrow \alpha.\, \varphi)^* \overset{\text{def}}{=} \varphi^*[\tilde{\alpha} \leftarrow w^*]$,
- $(@_n \varphi)^* \overset{\text{def}}{=} (\varphi^*[w_0^* \leftarrow \tilde{n}])[w^* \leftarrow \tilde{n}]$,
- $n^* \overset{\text{def}}{=} (w^* = \tilde{n})$, and
- $\alpha^* \overset{\text{def}}{=} (w^* = \tilde{\alpha})$.

Interpretation of $\tilde{n}$ in the corresponding first-order model is done in the obvious way, namely,

$$\tilde{n}^{\mathfrak{M}^*} = n^{\mathfrak{M}},$$

where $\mathfrak{M}$ is a modal model. By simple induction we again have for any t-sentence φ

$$\mathfrak{M} \models \varphi \iff \mathfrak{M}^* \models \varphi^*[w^* \leftarrow w_0^*].$$

By considering the translation of hybrid formulas into first-order ones, as long as no new formulas are involved (like interpolants), we can use the completeness, compactness and deduction theorems.

5.2 Craig's interoplation and Beth's definability theorems in QHL

Quantified Hybrid Logic repairs the failure for Beth's property by making Craig's interpolation theorem holds. To prove Craig's interpolation theorem we will need the following fact.

Lemma 5.1. *Let $n_1, \ldots, n_l$ be nominals. Let φ and $\theta(n_1, \ldots, n_l)$ be quantified hybrid formulas such none of the n_i's occur in φ. Let $\theta(\alpha_1, \ldots, \alpha_l)$ be $\theta(n_1, \ldots, n_l)$ in which each n_i is replaced by α_i. Then,*

(i) if $\vDash \varphi \rightarrow \theta(n_1, \ldots, n_l)$, then $\vDash \varphi \rightarrow \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)$, and

(ii) if $\vDash \theta(n_1, \ldots, n_l) \rightarrow \varphi$, then $\vDash \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l) \rightarrow \varphi$.

Proof. For (i), let $\vDash \varphi \rightarrow \theta(n_1, \ldots, n_l)$ and the n_i such that they do not occur in φ. We want to show that $\vDash \varphi \rightarrow \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)$, that is for any structure $\mathfrak{S} : \langle W, R, D, m \rangle$ in the vocabulary of $\{\varphi, \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)\}$ and any $w \in W$, $\langle \mathfrak{S}, w \rangle \vDash \varphi \rightarrow \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)$. Assume that $\langle \mathfrak{S}, w \rangle \vDash \varphi$. We can expand the structure $\mathfrak{S}$ into a structure $\mathfrak{S}'$ with nominals $n_1, \ldots, n_l$ such that $n_i^{\mathfrak{S}'} := w$ for $i = 1, \ldots, l$. Since φ do not contain the n_i, $\langle \mathfrak{S}', w \rangle \vDash \varphi$. Therefore $\langle \mathfrak{S}', w \rangle \vDash \theta(n_1, \ldots, n_l)$ which is equivalent to $\langle \mathfrak{S}, w \rangle \vDash \varphi \rightarrow \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)$.

For (ii), if $\nvDash \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l) \rightarrow \varphi$, then there exists a structure $\mathfrak{S} := \langle W, R, D, m \rangle$, a $w \in W$ and a valuation k such that $\langle \mathfrak{S}, w \rangle \vDash \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)[k]$

but $\langle \mathfrak{S}, w\rangle \not\models \varphi$. Change (or expand) $\mathfrak{S}$ into $\mathfrak{S}'$ by only changing the valuation of the nominals $n_1, \ldots, n_l$ such that for all n_i, $n_i^{\mathfrak{S}'} = w$. Then $\langle \mathfrak{S}', w\rangle \models \theta(n_1, \ldots, n_l)[k]$, and as the n_i do not occur in φ, still $\langle \mathfrak{S}', w\rangle \not\models [k]$. Thus $\langle \mathfrak{S}', w\rangle \not\models \theta(n_1, \ldots, n_l) \to \varphi[k]$. □

Remark 5.2. By the deduction theorem, if φ is a sentence, then we have

(i) if $\varphi \models \theta(n_1, \ldots, n_l)$, then $\varphi \models \downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l)$, and

(ii) if $\theta(n_1, \ldots, n_l) \models \varphi$, then $\downarrow \alpha_1 \ldots \downarrow \alpha_l . \theta(\alpha_1, \ldots, \alpha_l) \models \varphi$.

Theorem 5.3 (Craig's interpolation theorem)**.** *Let φ be a t_1-sentence and ψ be a t_2-sentence. If $\varphi \models \psi$ then there exists a $t_1 \cap t_2$-sentence θ such that $\varphi \models \theta$ and $\theta \models \varphi$.*

Proof. We will follow closely the proof given for theorem 3.2.

Let φ and ψ be quantified hybrid sentences. Without loss of generality, we may assume that φ and ψ are boolean combination of sentences (such sentences are called closed sentences in [3]) of the form $@_n\theta$ for $n \in \mathsf{NOM}$. In fact, suppose that φ and ψ are just sentences. Let n be a nominal not occurring in φ and ψ. If $\varphi \models \psi$, then also $@_n\varphi \models @_n\psi$. Let *theta* be an interpolant of $@_n\varphi$ and $@_n\psi$. As n does not occur in φ nor in ψ, $\downarrow \alpha . \theta[n \leftarrow \alpha]$ is an interpolant of φ and ψ by lemma 5.1. We want to deal only with closed sentences because their first order translations are also sentences. We can then either reason on the sentences as hybrid sentences or first order sentences. Using the first order perspective, we can apply the basic results on completeness and compactness and the deduction theorem.

Assume that φ and ψ have no interpolant θ. Then, we will derive a contradiction by showing that $\varphi \wedge \neg\psi$ has a model.

Let $t_0 = t_1 \cap t_2$. Let C be a countable infinite set of constant symbols not occurring in $t_1 \cup t_2$. Let N be a countable infinite set of nominals not occurring in $t_1 \cup t_2$. Let $t'_i = t_i \cup C \cup N$, for $i = 0, 1, 2$. Suppose that T is a t'_1-theory and U is a t'_2-theory.

Claim 5.3.1. *The theories $\{\varphi\}$ and $\{\neg\psi\}$ are inseparable.*

Proof. For the sake of contradiction, assume that there exists a t'_0-sentence θ separating $\{\varphi\}$ and $\{\neg\psi\}$. Then we have $\varphi \vDash \theta$ and $\neg\psi \vDash \neg\theta$ or equivalently $\theta \vDash \psi$. We may assume that θ has the form $\theta'(c_1, \ldots, c_l, n_1, \ldots, n_{l'})$, where $c_i \in C$ for $i = 1, \ldots, n$, $n_i \in N$ for $i = 1, \ldots, l'$. Therefore, by lemma 5.1, $\varphi \vDash \downarrow \alpha_1 \ldots \downarrow \alpha_{l'}.\forall x_1 \ldots \forall x_l \theta'(x_1, \ldots, x_l, \alpha_1, \ldots, \alpha_{l'})$ and $\downarrow \alpha_1 \ldots \downarrow \alpha_{l'}.\forall x_1 \ldots \forall x_l \theta'(x_1, \ldots, x_l, \alpha_1, \ldots, \alpha_{l'}) \vDash \psi$, contradicting the fact that φ and ψ have no interpolant. □

Let $\varphi_i, i < \omega$ and $\psi_i, i < \omega$ be enumerations of all closed t_1-sentences in and all closed t_2-sentences, respectively. We will construct two increasing sequences of theories (containing only closed sentences)

$$\{\varphi\} = T_0 \subseteq T_1 \subseteq T_2 \ldots$$

$$\{\neg\psi\} = U_0 \subseteq U_1 \subseteq U_2 \ldots$$

in the language of t'_1 and t'_2, respectively, such that for all $i < \omega$:

1. T_i and U_i are inseparable.

2. (a) if $T_i \cup \{\varphi_i\}$ and U_i are inseparable then $\varphi_i \in T_{i+1}$, and

   (b) if T_{i+1} and $U_i \cup \{\psi_i\}$ are inseparable then $\psi_i \in U_{i+1}$,

3. (a) if φ_i has the form $@_n \exists x \sigma(x)$ and $\varphi_{i+1} \in T_i$ then $@_n \sigma(c) \in T_{i+1}$ for some $c \in C$, and

   (b) if ψ_i has the form $@_n \exists x \sigma(x)$ and $\psi_i \in U_{i+1}$then $@_n \sigma(c) \in U_{i+1}$ for some $c \in C$,

4. (a) if φ_i has the form $@_n \Diamond \sigma$ and $\varphi_{i+1} \in T_i$ then $@_n \Diamond n' \wedge @_{n'} \sigma \in T_{i+1}$ for some $n' \in N$, and

(b) if ψ_i has the form $@_n\Diamond\sigma$ and $\psi_{i+1} \in U_{i+1}$, then $@_n\Diamond n' \wedge @_{n'}\sigma \in U_{i+1}$ for some $n' \in N$.

Given T_i and U_i, T_{i+1} and U_{i+1} are again constructed in the obvious way.

For (3) and (4), the constant c and the nominal n' are chosen such that they did not occur in T_i, U_i, φ_i or ψ_i. In that way, inseparability is preserved. We need to be worried only with (4). In fact, if $\mathfrak{S}$ is a modal structure such that $\langle \mathfrak{S}, w_0 \rangle \vDash @_n\Diamond\varphi$, then one can expand $\mathfrak{S}$ into $\mathfrak{S}'$ with a vocabulary containing the new nominal n' such that $\langle \mathfrak{S}', w_0 \rangle \vDash @_n\Diamond n' \wedge @_{n'}\varphi$. This is the case since

$$\langle \mathfrak{S}', w_0 \rangle \vDash @_n\Diamond n' \iff nRn'.$$

Let $T_\omega = \bigcup_{i<\omega} T_i$ and $U_\omega = \bigcup_{i<\omega} U_i$. Since every T_i and U_i are finite theories for $i < \omega$, by the Compactness theorem, it follows that T_ω and U_ω are inseparable.

We have the following claims whose proofs are exactly like in the first order case.

Claim 5.3.2. *The theories T_i and U_i are consistent for every $i \leq \omega$.*

Claim 5.3.3. *The theories T_ω and U_ω are maximal with respect to closed sentences.*

Claim 5.3.4. *The t_0'-theory $T_\omega \cap U_\omega$ is maximal consistent.*

Since T_ω is consistent, let $\mathfrak{N}_1$ be a t_1'-model such that $\mathfrak{N}_1 \vDash T_\omega$. Let n be a nominal in t_1. For any constant symbol $e \in t_1$, $\mathfrak{N}_1 \vDash @_n\exists x(e = x)$. By maximality of T_ω, $@_n\exists x(e = x) \in T_\omega$. Using (3), we can then construct a first-order submodel $\mathfrak{M}_1^*$ of $\mathfrak{N}_1^*$ such that

- $\mathcal{U}(\mathfrak{M}_1^*) = C_1^* \cup N_1^*$, where $C_1^* = \left\{c^{\mathfrak{N}_1^*} : c \in C\right\}$ and $N_1^* = \{\bar{n} : n \text{ nominals in } t_1'\}$. If we use the notation we have adopted for standard translation $\mathfrak{M}^* := \langle W \cup \bar{D}, m^* \rangle$ of a modal model $\mathfrak{M}$, then here $\tilde{C}$ plays the role of $\bar{D}$ and $\tilde{N}$ plays the role of W,
- $e^{\mathfrak{M}_1^*} = e^{\mathfrak{N}_1^*}$ for every constant symbol in $(t_1')^*$, and

- interpretations of relation symbols in $\mathfrak{M}_1^*$ are their interpretations in $\mathfrak{N}_1^*$ restricted to $\mathcal{U}(\mathfrak{M}_1^*)$.

Using Tarski-Vaught criterion and (3) and (4), we have the following claim, where T_ω^* is the theory resulting from translating every sentence of T_ω.

Claim 5.3.5. *We have $\mathfrak{M}_1^* \preccurlyeq \mathfrak{N}_1^*$, and in particular $\mathfrak{M}_1^* \models T_\omega^*$.*

In the same way, let $\mathfrak{N}_2$ be a t_2'-structure such that $\mathfrak{N}_2 \models U_\omega$. We can also construct a first order elementary substructure $\mathfrak{M}_2^*$ of $\mathfrak{N}_2^*$ such that $\mathcal{U}(\mathfrak{M}_2^*) = C_2^* \cup N_2^*$, where $C_2^* = \left\{c^{\mathfrak{N}_2^*} : c \in C\right\}$ and $N_2^* = \{\bar{n} : n \text{ nominals in } t_2'\}$. In particular, $\mathfrak{M}_2^* \models U_\omega^*$.

Like in the first order case,

Claim 5.3.6. *We have $\mathfrak{M}_1^* \restriction (t_0')^* \cong \mathfrak{M}_2^* \restriction (t_0')^*$.*

Based on that isomorphism we can extend the model for T_ω^* to a model for U_ω^* as well. Since $\varphi^* \in T_\omega^*$ and $\neg\psi \in U_\omega^*$, we constructed a model for $\varphi^* \wedge \neg\psi^*$. But then $\varphi \wedge \neg\psi$ has a model also. This ends the proof of Craig's interpolation theorem for QHL. □

From Craig's interpolation theorem, using the same definition of implicit and explicit definition in Definition 3.3, we have the Beth's definability theorem for QHL.

Theorem 5.4 (Beth definability theorem)**.** *Let $\Sigma(r)$ be a $t \cup \{r\}$-theory for some vocabulary t and $r \notin t$. Then, $\Sigma(r)$ defines R explicitly if and only if it defines R implicitly.*

5.3 Discussion

Let us go back to the counterexample given in Theorem 4.4. We saw that $\langle \mathfrak{S}, w_0 \rangle \models p$ if and only if

there exists $w \in W$ such that $\bar{r}_w \cap \bar{r}_{w_0} = \emptyset$. (4.2.1)

The condition (4.2.1) cannot be expressed in Quantified Modal Logic. In Quantified Hybrid Logic, we have in any structure $\mathfrak{S}$ (4.2.1) if and only if

$$\langle \mathfrak{S}, w_0 \rangle \vDash \downarrow \alpha. \Diamond \forall x (rx \to @_\alpha \neg rx).$$

Here, we see that naming the current state of evaluation using $\downarrow \alpha$ and referring back to it with $@_\alpha$ enables us to express (4.2.1) in the language.

Now we would like to highlight the steps needed to achieve Beth property.

Firstly, the proof relies heavily on the completeness and compactness of First Order Logic.

Secondly, in order to prove that $\{\varphi\}$ and $\{\neg\psi\}$ are inseparable, the $\downarrow$-binder was important. In the proof of Craig's interpolation theorem for first order logic, we mainly rely on the fact that we can introduce new constants to name objects and that we can go back to the original language by using quantifiers. In modal logic, naming worlds is impossible. Naming worlds is achieved by the use of nominals in hybrid logic. However, in order to stay in the original language, the use of some quantifier is required. One can introduce the use of $\forall$ to quantify over worlds. Unfortunately, the use such quantifier will lose the locality of modal logic: only reachable worlds are relevant for semantic evaluations. The use of $\downarrow$-binder keeps this local property of modal logic and allows us to go back to the original language.

Thirdly, for the theories T^*_ω and U^*_ω, there is a witness for each existential quantifier $\exists x^*$ such that $\exists x^* \sigma(x^*)$ is in the theory. If we translate a modal formula into a first-order one, then x^* can refer to an individual or a world. The first case is dealt by introducing a new constant c. The second case is dealt by introducing a new nominal n' in the hybrid language. Such existential formula can then be witnessed in the hybrid language by the closed sentence $@_n \Diamond n' \wedge @_{n'} \sigma$.

Bibliography

[1] H. Andréka, J. X. Madarász, and I. Németi, *Definability of New Universes in Many-sorted Logic*, Preprint (2001), It will soon be available from the home page of the Rényi Institute.

[2] H. Andréka, I. Németi, and I. Sain, *Universal algebraic logic (dedicated to the unity of science)*, Studies in Universal Logic, Birkhäuser (Basel–Boston–Stuttgart), 2008.

[3] C. Areces, P. Blackburn, and M. Marx, *Repairing the interpolation theorem in quantified modal logic*, Annals of Pure and Applied Logic **124** (2003), no. 1-3, 287–299.

[4] C. H. Bergman, R. D. Maddux, and D. L. Pigozzi (eds.), *Algebraic logic and universal algebra in computer science. lecture notes in computer science*, vol. 425, Springer–Verlag, Berlin, 1990, xi+292 p.

[5] E.W. Beth, *On Padoa's method in the theory of definition*, Nederl. Akad. Wetensch. Proc. Ser. A, vol. 56, 1953, pp. 330–339.

[6] W. J. Blok and D. L. Pigozzi, *Algebraizable logics*, Memoirs Amer. Math. Soc. **77, 396** (1989), vi+78.

[7] C.C. Chang and H.J. Keisler, *Model theory*, North Holland, 1990.

[8] K. Fine, *Model theory for modal logic part IThe De re/de dicto distinction*, Journal of philosophical logic **7** (1978), no. 1, 125–156.

[9] ———, *Failures of the interpolation lemma in quantified modal logic*, Journal of Symbolic Logic **44** (1979), no. 2, 201–206.

[10] L. Henkin, J. D. Monk, and A. Tarski, *Cylindric Algebras Parts I and II*, North–Holland, Amsterdam, 1971 and 1985.

[11] L. Henkin, J. D. Monk, A. Tarski, H. Andréka, and I. Németi, *Cylindric Set Algebras*, Lecture Notes in Mathematics, vol. 883, Springer-Verlag, Berlin, 1981, vi+323 pp.

[12] J. D. Monk, *An Introduction to Cylindric Set Algebras*, Logic Journal of the IGPL **8** (July 2000), no. 4, 449–494, Electronically available as: http://www.jigpal.oupjournals.org.

[13] H. Reichenbach, *The Theory of Relativity and A priori Knowledge*, University of California Press, 1960, Translated by M. Reichenbach from the original German edition published in 1920.

[14] I. Sain, *Beth's and Craig's properties via epimorphisms and amalgamation in algebraic logic*, In [4] (1990), 209–226.

[15] A. Tarski, *Some methodological investigations on the definability of concepts (in germain)*, Collected Papers, Volumes 1–4 (Steven R. Givant and Ralph N. McKenzie, eds.), Birkhäuser (Basel–Boston–Stuttgart), 1986, The original Polish version contains more detail: in Revue Philosophique, Vol.37, 1934, pp. 637–639.