Thesis for the Degree of Licentiate of Philosophy

Satisfaction classes in nonstandard models of first-order arithmetic

Fredrik Engström

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg, Sweden

Göteborg, August 2002

Satisfaction classes in nonstandard models of first-order arithmetic
FREDRIK ENGSTRÖM

ISSN 0347-2809/NO 2002:24
Department of Mathematics
Chalmers University of Technology and Göteborg University
412 96 Göteborg
Sweden
Telephone +46 (0)31-772 1000

Matematiskt centrum
Göteborg, Sweden 2002

Abstract

A satisfaction class is a set of nonstandard sentences respecting Tarski's truth definition. We are mainly interested in full satisfaction class, i.e., satisfaction classes which decides all nonstandard sentences. Kotlarski, Krajewski and Lachlan proved in 1981 that a countable model of `PA` admits a satisfaction class if and only if it is recursively saturated. A proof of this fact is presented in detail in such a way that it is adaptable to a language with function symbols. The idea that a satisfaction class can only see finitely deep in a formula is extended to terms. The definition gives rise to new notions of valuations of nonstandard terms; these are investigated. The notion of a *free* satisfaction class is introduced, it is a satisfaction class free of existential assumptions on nonstandard terms.

It is well known that pathologies arise in some satisfaction classes. Ideas of how to remove those are presented in the last chapter. This is done mainly by adding inference rules to $\mathfrak{M}$-logic. The consistency of many of these extensions is left as an open question.

Keywords: Satisfaction classes, Models of arithmetic.

AMS 2000 Subject Classification: 03C50, 03C62, 03H15

Acknowledgement

My supervisor, Richard W. Kaye, deserves all my thanks for letting me come to Birmingham under his inspiring supervision, and, of course, for showing me 'the way' the numerous times I got lost in the jungle of logic. I would also like to thank Thierry Coquand, who is my supervisor in Göteborg, and Jan Smith, who introduced me to logic, for their support which made it possible for me to work in models of PA.

Also, I would like to thank all at Mathematical Sciences at Göteborg University and Chalmers University of Technology, and the School of Mathematics and Statistics at the University of Birmingham.

Thank you!

Contents

1

Introduction

By the work of Skolem we know there are nonstandard models of Peano Arithmetic ($\mathtt{PA}$) and we know how to arithmetise logic inside $\mathtt{PA}$ due to Gödel. It is also easy to see that in any such nonstandard model there are nonstandard elements which the model thinks are sentences. By Tarski's truth definition we also know what it means for a standard sentence to be true. The obvious question is:

When is a nonstandard sentence true?

Given a nonstandard model $\mathfrak{M}$ of $\mathtt{PA}$ a satisfaction class is a (non-definable) predicate which, in a special sense, is a truth definition for nonstandard sentences, i.e., it respects Tarski's truth definition. More formally, a satisfaction class is a set Σ of standard and nonstandard sentences, extending the elementary diagram of $\mathfrak{M}$, such that

$$\begin{array}{rll} \neg\varphi \in \Sigma & \text{iff} & \varphi \notin \Sigma, \\ \varphi \vee \psi \in \Sigma & \text{iff} & \varphi \in \Sigma \text{ or } \psi \in \Sigma, \qquad \text{and} \\ \exists \mathsf{v}_i\, \gamma \in \Sigma & \text{iff} & \text{there exists } a \in \mathfrak{M} \text{ such that } \gamma[a/\mathsf{v}_i] \in \Sigma, \end{array}$$

for any nonstandard sentences φ, ψ and $\exists \mathsf{v}_i \gamma$. It is not obvious that we can construct such a set, in fact we cannot always do this, depending on the saturation of $\mathfrak{M}$.

This chapter is a short introduction to the subject with some motivation of the study and a historical overview. Next chapter is intended as a review of the prerequisites for this thesis; the results are given without proofs. In Chapter 3 we have rewritten the construction of satisfaction classes in a new style, the language includes functions, as opposed to [KKL81], and the terms are treated as they should be treated, i.e., satisfaction classes can only "look" finitely deep into them, as opposed to [Kay91].

In Chapter 4 we study some alternative definitions of satisfaction classes which all are weaker in the sense that the sentence $\exists x(t = x)$ does not have to be true.

We call these satisfaction classes 'free' since they are free of existential assumptions on nonstandard terms.

The last chapter is devoted to pathological examples that arise in satisfaction classes and how to remove them. We introduce satisfaction classes closed under propositional proofs and even stronger notions.

All structures studied in this thesis are models of PA in the language $\mathscr{L}_A$ with symbols $\{ \mathsf{S}, +, \cdot, 0 \}$.

1.1 Motivation

From the philosophical point of view there is an obvious motivation for the study of satisfaction classes. We know what truth is for standard formulas by the truth definition of Tarski and with the arithmetisation of logic we have the notion of nonstandard formulas. The obvious question is then, what does it mean for a nonstandard formula to be true? The study of satisfaction classes is an attempt to answer this question. In fact, satisfaction classes can be seen as nonstandard models in the same way as complete consistent Henkin theories represent their Henkin models. There are $2^{\aleph_0}$ satisfaction classes in a countable recursively saturated model of PA, which implies that the structure of satisfaction classes is rich in some sense[1] and we can see the study of satisfaction classes as a nonstandard model theory.

In [Smi84], Smith characterises the recursively saturated models of PA in terms of satisfaction classes, he constructs a Σ^1_1 formula characterising them. He also Δ^1_2 characterises resplendent models in terms of satisfaction classes. This gives us some idea that the satisfaction classes are more important, in a mathematical sense, then just as truth definitions. For mathematicians there are a lot of questions to be answered by this study. For example, one of my motivations for this thesis has been to find characterisations of other interesting model theoretic properties, such as saturation properties stronger than recursive saturation.

The study could also be seen as an example of how to work with ill founded objects, there may even be applications to computer science.

The main reason for the existence of this thesis is to show that there are many unstudied related notions of satisfaction classes. Some of them arise naturally when we add function symbols in the language and others when we try to remove certain "pathologies."

1.2 Historical background

In 1963 Abraham Robinson published a paper, 'On languages which are based on nonstandard arithmetic' [Rob63], where he discusses syntax and semantics for nonstandard languages. This is, as far as I know, the first time nonstandard

[1] In fact, the set of satisfaction classes is dense in the Stone space of the nonstandard Lindenbaum algebra, see [Smi84].

languages are defined and investigated explicitly. He does not use the word satisfaction class, but he gives two different examples of semantics for nonstandard languages. He calls them the internal and the external truth definitions. The external one, defined by the help of Skolem operators, is defined only for formulas with finite 'Robinson-rank' (which is a complexity measure on nonstandard formulas), therefore it is not a full satisfaction class.

The internal truth definition is defined as Σ if

$$(\mathbb{N}, \Sigma_0) \prec \langle \mathfrak{M}, \Sigma \rangle \,,^{2}$$

where $\mathbb{N}$ is the standard model of arithmetic and Σ_0 is the standard truth definition, i.e.,

$$\Sigma_0 = \{ \varphi \,|\, \varphi \text{ is a standard sentence in } \mathscr{L}_A \text{ and } \mathbb{N} \vDash \varphi \} \,.$$

He proves that the external and the internal truth definitions do not coincide and leaves it, more or less, there.

Later Krajewski [Kra76] returns to the question of the semantics of nonstandard languages. He defines satisfaction classes (even though his definition is rather weak) and investigates some related notions. He also proves that for some models of cardinality λ there exists 2^λ full satisfaction classes. He does not mention the question of which models admit satisfaction classes, he only proves that some specific models do.

In [KKL81] and [Lac81] this question is answered. In the first paper it is shown that if a countable model is recursively saturated then it admits a satisfaction class. This is done by using a version of ω-logic (the idea is due to Jeff Paris) and the result is not surprising. What is more surprising is the result in the second paper by Lachlan in which he shows that if a model (of any cardinality) admits a satisfaction class then it is recursively saturated. He proves this using a sort of overspill he gets from the satisfaction class (proving the result with induction in the language $\mathscr{L}_A \cup \{ \Sigma \}$ is easy).

In [Smi84] Smith strengthens the results in the two papers above. He shows that any resplendent model has a satisfaction class and is able to find a Δ^1_2 characterisation of resplendency. He also formulates Lachlan's proof in a syntactical way which makes it possible to find a Σ^1_1 formula characterising recursive saturation.

Other important contributions to the study of satisfaction classes are [Kot85], [KR90a] and [KR90b]. These papers discuss the question of when the structure $\langle \mathfrak{M}, \Sigma \rangle$ satisfies either induction over Σ_k formulas or full induction. In the first paper Kotlarski shows that the satisfaction classes satisfying Δ_0 induction are precisely those closed under nonstandard proofs of first-order logic and including all nonstandard instances of the axiom of induction. It is easy to see that if a model admits such a satisfaction class then it satisfies the scheme of reflection:

$$\text{`}\mathtt{PA} \vdash \varphi\text{'} \rightarrow \varphi.$$

[2] In fact Robinson's notion is a bit different, but for the purpose of this survey we can think of internal truth defined in this way.

The definitions of satisfaction classes in the preceding works are all using a relational language, i.e., the language $0, 1, \sigma, \pi$ where σ and π are ternary relational symbols, supposed to express addition and multiplication. The theory $\mathtt{PA}$ has to be extended by some axioms expressing that these relational symbols are in fact functions. In [Kay91] Kaye investigates the case when $\mathtt{PA}$ is expressed in the language with symbols: $+$, $\cdot$, $<$, 0 and 1, where $+$ and $\cdot$ are binary function symbols. Throughout this thesis we will be using the language $\mathscr{L}_A$ which has one unary function symbol S, two binary function symbols $+$ and $\cdot$, and one constant symbol 0 (and the equality predicate $=$).

1.3 Theorems, Propositions, Lemmas, Corollaries and Porisms

We will use the term 'theorem' sparsely, it is used to put extra emphasise on an important result. Propositions are the results which have a value on their own, and lemmas (or lemmata) are results which help us to prove propositions or theorems. Corollaries are simple consequences of propositions or theorems (and in some rare occasions of lemmas), but the, somewhat, unusual term porism is used for simple consequences of a *proof* of a proposition, theorem or lemma. It could for example be a simple generalisation of a proposition which you get by some minor modifications of the presented proof of the proposition. The term was used by Euclid, but he used it with a somewhat different meaning, which is not entirely known, see [Twe00] for more information.

2

Prerequisites

To be able to read this thesis the reader needs some background knowledge of first-order logic, [Men97] is more than enough, and also some knowledge of first-order arithmetic (Peano Arithmetic), especially its model theory. A good general reference for this is [Kay91]. In this chapter we will review some of the material, omitting the proofs of the results.

2.1 Peano Arithmetic

Throughout this thesis $\mathfrak{M}$ will be a structure in the language $\mathscr{L}_A$ with symbols S, $+$, $\cdot$ and 0, where S is a unary function, $+$ and $\cdot$ are binary function symbols and 0 is a constant symbol. We will as usual write terms and equalities in the more convenient way by using infix notation, e.g., instead of writing $+\big(t_1, t_2\big)$ we will write $t_1 + t_2$.

We denote interpretations as usually; the interpretation of S in $\mathfrak{M}$ is denoted by $\mathsf{S}^{\mathfrak{M}}$, the interpretation of $+$ is denoted $+^{\mathfrak{M}}$, and so on.

We will consider formulas built up from the logical connectives $\neg$ and $\vee$ and the quantifier $\exists$. The other connectives and quantifiers are considered to be abbreviations in the usual way, $\varphi \wedge \psi$ is *defined* to be $\neg\big(\neg\varphi \vee \neg\psi\big)$, $\varphi \rightarrow \psi$ is $\neg\varphi \vee \psi$, $\varphi \leftrightarrow \psi$ is $\big(\varphi \rightarrow \psi\big) \wedge \big(\psi \rightarrow \varphi\big)$ and $\forall \mathsf{v}_i\, \varphi$ is $\neg\exists \mathsf{v}_i\, \neg\varphi$. We also define exclusive or $\varphi \dot{\vee} \psi$ as $(\varphi \vee \psi) \wedge \neg(\varphi \wedge \psi)$. Later it will be important to note that all these abbreviations are of constant depth, i.e., the depth only increases a constant number when replacing the definiendum with the definiens. For example, the depth of $\neg\big(\neg\varphi \vee \neg\psi\big)$ is always two more than that of $\varphi \wedge \psi$.

The variables in the language are $\mathsf{v}_0, \mathsf{v}_1, \ldots$. Sometimes we will be a bit sloppy in the notation and use $x, y, w, \ldots$ as names for variables. If φ is a formula then

$$\varphi[t_1, \ldots, t_k/\mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_k}]$$

will denote the formula you get by substituting all free occurrences of v_{i_l} with the term t_l, we will always assume that v_{i_l} is free for t_l. Sometimes, when it will not cause confusion, we will write $\varphi(t)$ to mean the formula obtained from φ be replacing all the occurrences of the free variable *under consideration* by the term t, i.e., $\varphi[t/\mathsf{v}_i]$ where v_i is the variable under consideration. In short, we will adopt all the usual notations and abbreviations used in the literature.

Let $\mathscr{L}_{\mathfrak{M}}$ be the language $\mathscr{L}_A \cup \{\, \mathsf{c}_a \mid a \in \mathfrak{M} \,\}$ (0 and $\mathsf{c}_{0^{\mathfrak{M}}}$ will be regarded as the same symbol; this is to simplify some definitions below) and let $\mathrm{ElDiag}(\mathfrak{M})$ be the theory of the structure $\mathfrak{M}$ in the language $\mathscr{L}_{\mathfrak{M}}$, i.e., all $\mathscr{L}_{\mathfrak{M}}$-formulas true in $\mathfrak{M}^+$, where $\mathfrak{M}^+$ is the expanded model of $\mathfrak{M}$ which interprets each symbol c_a as a. Mostly we will not distinguish between the two structures $\mathfrak{M}$ and $\mathfrak{M}^+$, hoping this will not cause any confusion for the reader. Sometimes, mostly when dealing with standard formulas and terms, we will identify the element $a \in \mathfrak{M}$ with the constant symbol c_a.

$\mathrm{Diag}(\mathfrak{M})$ is the set of all true standard atomic and negated atomic formulas in the language $\mathscr{L}_{\mathfrak{M}}$, so $\mathrm{Diag}(\mathfrak{M}) \subsetneq \mathrm{ElDiag}(\mathfrak{M})$.

$\mathtt{PA}^-$ is the theory with the universal closures of

$$
\begin{gathered}
\mathsf{S}(x) = \mathsf{S}(y) \to x = y, \\
\mathsf{S}(x) \neq 0, \\
x \neq 0 \to \exists y\big(\mathsf{S}(y) = x\big), \\
x + 0 = x, \\
x + \mathsf{S}(y) = \mathsf{S}(x + y), \\
x \cdot 0 = 0 \quad \text{and} \\
x \cdot \mathsf{S}(y) = x \cdot y + x
\end{gathered}
$$

as axioms. If we add the axiom scheme of induction:

$$
\forall \bar{x}\Big(\varphi(0, \bar{x}) \land \forall y\big(\varphi(y, \bar{x}) \to \varphi(\mathsf{S}(y), \bar{x})\big) \to \forall y\, \varphi(y, \bar{x})\Big)
$$

for all $\mathscr{L}_A$-formulas φ, we get Peano Arithmetic or $\mathtt{PA}$ for short.

The symbol $\mathfrak{M}$ will always be assumed to be a structure, in a language extending $\mathscr{L}_A$, satisfying $\mathtt{PA}$ and not isomorphic to the standard model $\mathbb{N}$ of $\mathtt{PA}$, i.e., $\mathfrak{M}$ will be assumed to be a nonstandard model of $\mathtt{PA}$.

The predicate $x < y$ is defined as

$$
\exists z\big(z \neq 0 \land x + z = y\big).
$$

Once again, it will be important later that this definition is of constant depth, i.e., $\exists z(z \neq 0 \land t + z = r)$ is of depth at most four more than the depth of $t < r$. We also define the function $x - 1$ by the following equation

$$
x - 1 =_{\mathrm{df}} \ (\mu z)\big(x = 0 \land z = 0\big) \lor \big(x \neq 0 \land \mathsf{S}(z) = x\big),
$$

where $(\mu x)\varphi(x)$ means 'the least x such that $\varphi(x)$.'

We will also identify the smallest initial segment of $\mathfrak{M}$ (i.e., the smallest non-empty subset of $\mathfrak{M}$ closed under successor and less then) with the standard model $\mathbb{N}$, i.e., we assume that $\mathbb{N} \subseteq \mathfrak{M}$. We will reserve the symbol $\mathbb{N}$ to denote the standard model and ω to denote the first infinite ordinal, which is the same as the domain of $\mathbb{N}$.

If $\varphi(x)$ is a formula with a free variable we will write $\varphi(\mathfrak{M})$ for the set of elements in $\mathfrak{M}$ satisfying $\varphi(x)$, i.e.,

$$\{\, a \in \mathfrak{M} \mid \mathfrak{M} \models \varphi(a) \,\} .$$

Given $a \in \mathfrak{M}$ we define $I_{<a}$ to be the initial segment

$$\{\, x \in \mathfrak{M} \mid x < a \,\} .$$

2.2 Coding

We will assume a notion of finite sets and a definable predicate $x \in y$ (i.e., an $\mathscr{L}_A$-formula with two free variables) such that the universal closures of

$$\begin{gathered} x \in y \to x < y, \\ \forall w\big(w \in x \leftrightarrow w \in y\big) \to x = y, \\ \exists z \forall y\big(y \in z \leftrightarrow y = x\big), \\ \exists z \forall w\big(w \in z \leftrightarrow \big[w \in x \vee w \in y\big]\big), \\ \exists z \forall w\big(w \in z \leftrightarrow \big[w \in x \wedge w \in y\big]\big), \quad \text{and} \\ \exists z \forall w\big(w \in z \leftrightarrow \big[w \in x \wedge w \notin y\big]\big), \end{gathered}$$

are all provable in `PA`. The sets z (which all are unique by the second property) in the last three formulas will be denoted $\{\, x \,\}$, $x \cup y$, $x \cap y$ and $x \setminus y$ respectively. We will also write $\{\, x, y \,\}$, $\{\, x, y, z \,\}$, $\dots$ for $\{\, x \,\} \cup \{\, y \,\}$, $\{\, x \,\} \cup \{\, y \,\} \cup \{\, z \,\}$, $\dots$ respectively. The membership predicate can be defined by using the exponentiation function, we will not go into the details of this here; for a good reference see [HP98].

We also assume a notion of finite sequences, either derived from the notion of finite sets (see [HP98]) or by the Chinese remainder theorem (see [Kay91]). The predicate $(x)_y = z$ is assumed to be such that the universal closures of the following formulas are provable in `PA`,

$$\begin{gathered} \exists! z\, (x)_y = z, \\ (x)_y \leqslant x, \\ \exists y\, (y)_0 = x, \quad \text{and} \\ \exists w\big(\forall i{<}y\, (x)_i = (w)_i \wedge (w)_y = z\big). \end{gathered}$$

We define the following provable recursive functions (and constant)

$$\begin{aligned}
\mathrm{len}(x) &=_{\mathrm{df}} (x)_0,\\
[x]_y &=_{\mathrm{df}} (\mu z)\big(y < \mathrm{len}(x) \wedge (x)_{\mathsf{S}(y)} = z\big) \vee \big(y \geqslant \mathrm{len}(x) \wedge z = 0\big),\\
x \frown y &=_{\mathrm{df}} (\mu z)\, \mathrm{len}(x) + \mathrm{len}(y) = \mathrm{len}(z)\\
&\qquad\qquad \wedge \forall i {<} \mathrm{len}(x)\, [x]_i = [z]_i \wedge \forall i {<} \mathrm{len}(y)\, [y]_i = [z]_{\mathrm{len}(x)+i},\\
[] &=_{\mathrm{df}} (\mu x)\, \mathrm{len}(x) = 0,\\
[x] &=_{\mathrm{df}} (\mu y)\, \mathrm{len}(y) = 1 \wedge [y]_0 = x,\\
[x_0, \ldots, x_k] &=_{\mathrm{df}} [x_0] \frown \cdots \frown [x_k],\\
x \restriction y &=_{\mathrm{df}} (\mu z)\, \mathrm{len}(z) = y \wedge \forall i {<} y\, [z]_i = [x]_i,
\end{aligned}$$

here $(\mu x)\varphi(x)$ means 'the least x such that $\varphi(x)$.'

2.3 Nonstandard languages

We need a Gödel numbering for the formulas and terms in the language $\mathscr{L}_{\mathfrak{M}}$. We might define the Gödel number for a formula φ, denoted $\ulcorner\varphi\urcorner$, to be (a code for) the sequence of the Gödel numbers of the symbols in φ, thus

$$\ulcorner \mathsf{S}(0) = \mathsf{v}_0 \urcorner = \Big[\ulcorner \mathsf{S}\urcorner, \ulcorner(\urcorner, \ulcorner 0\urcorner, \ulcorner)\urcorner, \ulcorner 0\urcorner, \ulcorner \mathsf{v}_0\urcorner\Big].$$

The exact definition we use for Gödel numbering is unimportant. But it will *not*, except in some special occasions, be assumed to be defined in this way; any numbering such that the properties below hold will work.

There are $\mathscr{L}_A$-formulas $\mathtt{Form}(x)$, $\mathtt{Sent}(x)$, $\mathtt{Term}(x)$ and $\mathtt{ClTerm}(x)$ coding, in $\mathtt{PA}$, the formulas, sentences, terms and closed terms of $\mathscr{L}_{\mathfrak{M}}$ respectively. Let $\mathtt{FV}$ be the function (defined in $\mathtt{PA}$) such that $\mathtt{FV}(\varphi)$ is (an element coding) the set of Gödel numbers of the free variables of φ. The precise construction of the formulas are not important, but some of the properties of them are. Those are listed below.

The notation $\ulcorner \mathsf{S}(x)\urcorner$ will be taken, when appropriate, to mean the *function* which takes a Gödel number of a term t and returns the Gödel number of the term $\mathsf{S}(t)$. By 'when appropriate' we mean that in some cases $\ulcorner \mathsf{S}(x)\urcorner$ means the Gödel number of the term $\mathsf{S}(x)$, but we will try to write $\ulcorner \mathsf{S}(\mathsf{v}_i)\urcorner$ in that case. Of course this also applies to for example $\ulcorner \exists \mathsf{v}_i\, x\urcorner$ which is a function taking a Gödel number of a formula φ and an i and returning the Gödel number of the formula $\exists \mathsf{v}_i\, \varphi$. With the assumption that a Gödel number is a sequence of (Gödel numbers of) symbols we get that

$$\mathtt{PA} \vdash \forall x \ulcorner \mathsf{S}(x)\urcorner = \Big[\ulcorner \mathsf{S}\urcorner, \ulcorner(\urcorner\Big] \frown x \frown \Big[\ulcorner)\urcorner\Big].$$

If x is not a Gödel number, or a Gödel number of "wrong type," then the functions can be defined to take the value 0.

In $\mathtt{PA}$ we can define the substitution function that takes the Gödel number of a formula/term, a term and a variable and returns the Gödel number of the formula/term we get by substituting all occurrences of the given variable with the given term. We denote this function $x[y/z]$.

That the universal closures of the following formulas are provable in $\mathtt{PA}$ tells us that all elements that should satisfy $\mathtt{Term}(x)$ does so.

$$\begin{gathered}
\mathtt{Term}(\ulcorner 0 \urcorner) \wedge \mathtt{Term}(\ulcorner \mathsf{c}_x \urcorner) \wedge \mathtt{Term}(\ulcorner \mathsf{v}_i \urcorner), \\
\mathtt{Term}(x) \rightarrow \mathtt{Term}(\ulcorner \mathsf{S}(x) \urcorner), \\
\mathtt{Term}(x) \wedge \mathtt{Term}(y) \rightarrow \mathtt{Term}(\ulcorner x + y \urcorner), \quad \text{and} \\
\mathtt{Term}(x) \wedge \mathtt{Term}(y) \rightarrow \mathtt{Term}(\ulcorner x \cdot y \urcorner).
\end{gathered}$$

The analogous formulas for $\mathtt{Form}(x)$ are:

$$\begin{gathered}
\mathtt{Term}(x) \wedge \mathtt{Term}(y) \rightarrow \mathtt{Form}(\ulcorner x = y \urcorner), \\
\mathtt{Form}(x) \rightarrow \mathtt{Form}(\ulcorner \neg x \urcorner), \\
\mathtt{Form}(x) \wedge \mathtt{Form}(y) \rightarrow \mathtt{Form}(\ulcorner x \vee y \urcorner) \quad \text{and} \\
\mathtt{Form}(x) \rightarrow \mathtt{Form}(\ulcorner \exists \mathsf{v}_i \, x \urcorner).
\end{gathered}$$

The next properties tell us that nothing other than what is supposed to satisfies $\mathtt{Term}(x)$. This is the inductive property of terms:

$$\begin{aligned}
\varphi(\ulcorner 0 \urcorner) \wedge \forall i \, \varphi(\ulcorner \mathsf{v}_i \urcorner) \wedge \forall x \, \varphi(\ulcorner \mathsf{c}_x \urcorner) \wedge \forall x, y \big[& \mathtt{Term}(x) \wedge \mathtt{Term}(y) \wedge \varphi(x) \\
& \wedge \varphi(y) \rightarrow \varphi(\ulcorner \mathsf{S}(x) \urcorner) \wedge \varphi(\ulcorner x + y \urcorner) \wedge \varphi(\ulcorner x \cdot y \urcorner) \big] \rightarrow \forall x \big(\mathtt{Term}(x) \rightarrow \varphi(x) \big)
\end{aligned}$$

for all $\mathscr{L}_{\mathfrak{M}}$-formulas $\varphi(x)$. The analogous property for $\mathtt{Form}(x)$:

$$\begin{aligned}
\forall x, y \big(\mathtt{Term}(x) \wedge \mathtt{Term}(y) \rightarrow \varphi(\ulcorner x = y \urcorner) \big) \wedge \forall x, y, i \big[& \mathtt{Form}(x) \wedge \mathtt{Form}(y) \\
& \wedge \varphi(x) \wedge \varphi(y) \rightarrow \varphi(\ulcorner \neg x \urcorner) \wedge \varphi(\ulcorner x \vee y \urcorner) \wedge \varphi(\ulcorner \exists \mathsf{v}_i x \urcorner) \big] \rightarrow \forall x \big(\mathtt{Form}(x) \rightarrow \varphi(x) \big)
\end{aligned}$$

for all $\mathscr{L}_{\mathfrak{M}}$-formulas $\varphi(x)$.

There are also some similar properties for $\mathtt{FV}(x)$.

$$\begin{aligned}
& \mathtt{FV}(\ulcorner 0 \urcorner) = 0 \wedge \forall x \; \mathtt{FV}(\ulcorner \mathsf{c}_x \urcorner) = 0, \\
& \forall i \, \mathtt{FV}(\ulcorner \mathsf{v}_i \urcorner) = \{ \ulcorner \mathsf{v}_i \urcorner \}, \\
& \forall x, y \big[\mathtt{Term}(x) \wedge \mathtt{Term}(y) \rightarrow \mathtt{FV}(\ulcorner \mathsf{S}(x) \urcorner) = \mathtt{FV}(x) \wedge \mathtt{FV}(\ulcorner x + y \urcorner) = \mathtt{FV}(\ulcorner x \cdot y \urcorner) \\
& \qquad = \mathtt{FV}(\ulcorner x = y \urcorner) = \mathtt{FV}(x) \cup \mathtt{FV}(y) \big], \quad \text{and} \\
& \forall x, y, i \big[\mathtt{Form}(x) \wedge \mathtt{Form}(y) \rightarrow \mathtt{FV}(\ulcorner \neg x \urcorner) = \mathtt{FV}(x) \wedge \mathtt{FV}(\ulcorner x \vee y \urcorner) = \mathtt{FV}(x) \cup \mathtt{FV}(y) \\
& \qquad \wedge \mathtt{FV}(\ulcorner \exists \mathsf{v}_i x \urcorner) = \mathtt{FV}(x) \setminus \{ \ulcorner \mathsf{v}_i \urcorner \} \big].
\end{aligned}$$

Last we have the defining properties of $\mathtt{ClTerm}(x)$ and $\mathtt{Sent}(x)$:

$$\begin{gathered}
\mathtt{ClTerm}(x) \leftrightarrow \mathtt{Term}(x) \wedge \mathtt{FV}(x) = 0 \quad \text{and} \\
\mathtt{Sent}(x) \leftrightarrow \mathtt{Form}(x) \wedge \mathtt{FV}(x) = 0.
\end{gathered}$$

Observe that these properties define the formulas $\mathtt{Term}(x)$, $\mathtt{Form}(x)$, $\mathtt{ClTerm}(x)$ and $\mathtt{Sent}(x)$ and the function $\mathtt{FV}(x)$ up to provable equivalence in $\mathtt{PA}$.

Two very important properties, which follows from the properties above, of $\mathtt{PA}$ is the following, usually called the 'unique readability property.' $\mathtt{PA}$ proves the following two sentences:

$$\begin{aligned}\forall x\Big(\mathtt{Term}(x) \to \big[&\exists! i\, x = \ulcorner \mathsf{v}_i \urcorner \dot\vee \exists! y\, x = \ulcorner \mathsf{c}_y \urcorner \dot\vee \exists! y\big(\mathtt{Term}(y) \wedge x = \ulcorner \mathsf{S}(y)\urcorner\big) \\ &\dot\vee \exists! y, z\big(\mathtt{Term}(y) \wedge \mathtt{Term}(z) \wedge x = \ulcorner y + z \urcorner\big) \\ &\dot\vee \exists! y, z\big(\mathtt{Term}(y) \wedge \mathtt{Term}(z) \wedge x = \ulcorner y \cdot z \urcorner\big)\big]\Big)\end{aligned}$$

and

$$\begin{aligned}\forall x\Big(\mathtt{Form}(x) \to \big[&\exists! y, z\big(\mathtt{Term}(y) \wedge \mathtt{Term}(z) \wedge x = \ulcorner z = y \urcorner\big) \\ &\dot\vee \exists! y\big(\mathtt{Form}(y) \wedge x = \ulcorner \neg y \urcorner\big) \dot\vee \exists! y, z\big(\mathtt{Form}(y) \wedge \mathtt{Form}(z) \wedge x = \ulcorner y \vee z \urcorner\big) \\ &\dot\vee \exists! y, i\big(\mathtt{Form}(y) \wedge x = \ulcorner \exists \mathsf{v}_i\, y \urcorner\big)\big]\Big).\end{aligned}$$

These properties make it possible to handle nonstandard languages. Let ${}^*\mathscr{L}_{\mathfrak{M}}$ be the nonstandard language which corresponds to $\mathscr{L}_{\mathfrak{M}}$, i.e., the "terms" of ${}^*\mathscr{L}_{\mathfrak{M}}$ are all $a \in \mathfrak{M}$ such that $\mathfrak{M} \models \mathtt{Term}(a)$ and the "formulas" are all $a \in \mathfrak{M} \models \mathtt{Form}(a)$, etc. The unique readability properties give us the possibility to handle these "terms" and "formulas" in much the same way as the standard ones, with the important exception that they need not be well-founded, e.g., $\neg\neg \ldots \neg 0 = 1$, where the dots represent a *nonstandard* number of negation signs, is a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula. Therefore, and this is very important, we do not have "external" induction on ${}^*\mathscr{L}_{\mathfrak{M}}$-terms and ${}^*\mathscr{L}_{\mathfrak{M}}$-formulas.[1]

2.4 Partial truth definitions

Due to Tarski's theorem on the undefinability of truth we cannot find an $\mathscr{L}_{\mathfrak{M}}$-formula φ such that

$$\mathtt{PA} \vdash \varphi(\ulcorner \psi \urcorner) \leftrightarrow \psi$$

for all $\mathscr{L}_{\mathfrak{M}}$-sentences ψ. What we can do is the following.

Theorem 2.1. *There is an $\mathscr{L}_A$-definable function* $\mathtt{val}$ *such that* $\mathtt{PA}$ *proves*

$$\mathtt{val}(\ulcorner t \urcorner) = t$$

for all closed $\mathscr{L}_{\mathfrak{M}}$-terms t, and

$$\begin{aligned}&\forall x\big(\mathtt{ClTerm}(x) \to \mathtt{val}(\ulcorner \mathsf{S}(x)\urcorner) = \mathsf{S}(\mathtt{val}(x))\big) \\ &\quad \wedge \forall x, y\big(\mathtt{ClTerm}(x) \wedge \mathtt{ClTerm}(y) \to \mathtt{val}(\ulcorner x + y \urcorner) = \mathtt{val}(x) + \mathtt{val}(y)\big) \\ &\qquad \wedge \forall x, y\big(\mathtt{ClTerm}(x) \wedge \mathtt{ClTerm}(y) \to \mathtt{val}(\ulcorner x \cdot y \urcorner) = \mathtt{val}(x) \cdot \mathtt{val}(y)\big).\end{aligned}$$

[1] We have induction *inside* the model, that is what the inductive property of $\mathtt{Form}(x)$ and $\mathtt{Term}(x)$ tells us.

Let Δ_k, Σ_k and Π_k be defined as usual, e.g., Σ_1 is the set of all formulas of the form $\exists\bar{x}\,\varphi(\bar{x})$ where $\varphi(\bar{x})$ is in Δ_0. There are $\mathscr{L}_A$-formulas coding these sets in $\mathtt{PA}$ in the usual sense. Let us write, for example, $x \in \Delta_0$ for the formula coding Δ_0 applied to the variable x.

Theorem 2.2. *There are $\mathscr{L}_A$-formulas $\mathtt{Tr}_\Gamma(x)$, where Γ is Δ_k, Σ_k or Π_k, such that* $\mathtt{PA}$ *proves*

$$\mathtt{Tr}_\Gamma(\ulcorner\psi\urcorner) \leftrightarrow \psi$$

for all sentences $\psi \in \Gamma$,

$$\forall x, y\big[\mathtt{ClTerm}(x) \wedge \mathtt{ClTerm}(y) \to \big(\mathtt{Tr}_{\Delta_0}(\ulcorner x = y\urcorner) \leftrightarrow \mathtt{val}(x) = \mathtt{val}(y)\big)\big],$$
$$\forall x\big[x \in \Delta_0 \wedge \mathtt{Sent}(x) \to \big(\mathtt{Tr}_{\Delta_0}(\ulcorner\neg x\urcorner) \leftrightarrow \neg\,\mathtt{Tr}_{\Delta_0}(x)\big)\big],$$
$$\forall x, y\big[x, y \in \Delta_0 \wedge \mathtt{Sent}(x) \wedge \mathtt{Sent}(y) \to \big(\mathtt{Tr}_{\Delta_0}(\ulcorner x \vee y\urcorner) \leftrightarrow \mathtt{Tr}_{\Delta_0}(x) \vee \mathtt{Tr}_{\Delta_0}(y)\big)\big],$$
$$\forall x, i, y\big[x \in \Delta_0 \wedge \mathtt{Sent}(\ulcorner\exists \mathsf{v}_i\, x\urcorner) \to \big(\mathtt{Tr}_{\Delta_0}(\ulcorner\exists \mathsf{v}_i{<}\mathsf{c}_y\, x\urcorner) \leftrightarrow \exists z{<}y\ \mathtt{Tr}_{\Delta_0}(\ulcorner x[\mathsf{c}_z/\mathsf{v}_i]\urcorner)\big)\big],$$

and also

$$\forall x, i\big[\ulcorner\exists \mathsf{v}_i\, x\urcorner \in \Gamma \wedge \mathtt{Sent}(\ulcorner\exists \mathsf{v}_i\, x\urcorner) \to \big(\mathtt{Tr}_\Gamma(\ulcorner\exists \mathsf{v}_i\, x\urcorner) \leftrightarrow \exists y\ \mathtt{Tr}_\Gamma(\ulcorner x[\mathsf{c}_y/\mathsf{v}_i]\urcorner)\big)\big], \quad \textit{and}$$
$$\forall x, i\big[\ulcorner\forall \mathsf{v}_i\, x\urcorner \in \Gamma \wedge \mathtt{Sent}(\ulcorner\forall \mathsf{v}_i\, x\urcorner) \to \big(\mathtt{Tr}_\Gamma(\ulcorner\forall \mathsf{v}_i\, x\urcorner) \leftrightarrow \forall y\ \mathtt{Tr}_\Gamma(\ulcorner x[\mathsf{c}_y/\mathsf{v}_i]\urcorner)\big)\big].$$

For explicit constructions see [Kay91].

2.5 Recursive saturation and resplendency

In this section $\mathfrak{M}$ can be any structure in any recursive language $\mathscr{L}$. Let $\mathscr{L}_\mathfrak{M}$ be the language which extends $\mathscr{L}$ with constant symbols naming all elements in $\mathfrak{M}$.

Given a theory T, a *type in T* is a countable set of formulas

$$t(\bar{x}) = \{\, \varphi_i(\bar{x}, \bar{a}) \mid i \in \omega \,\}$$

such that $\varphi_i(\bar{x}, \bar{y})$ are $\mathscr{L}$-formulas with finitely many free variables $\bar{x}$ and $\bar{y}$; $\bar{a}$ are parameters from $\mathfrak{M}$; and $T + t(\bar{\mathsf{c}})$, where $\bar{\mathsf{c}}$ are new constant symbols, is consistent. A type over a model $\mathfrak{M}$ is a type in the elementary diagram of the model, i.e., in ElDiag($\mathfrak{M}$). A type $t(\bar{x})$ is *recursive* if the set

$$\left\{ \ulcorner\varphi(\bar{x}, \bar{y})\urcorner \,\middle|\, \text{there exists } \bar{a} \in \mathfrak{M} \text{ such that } \varphi(\bar{x}, \bar{a}) \in t(\bar{x}) \right\} \subseteq \omega$$

is recursive. If $t(\bar{x})$ is a type then it is *realized* in $\mathfrak{M}$ if there are elements $\bar{m} \in \mathfrak{M}$ such that $\mathfrak{M} \models \varphi(\bar{m})$ for all $\varphi(\bar{x}) \in t(\bar{x})$.

Definition 2.3. $\mathfrak{M}$ is *recursively saturated* if all recursive types in $\mathfrak{M}$ are realized.

The following proposition says that all consistent theories have recursively saturated models of any cardinality.

Proposition 2.4. *For every $\mathfrak{M}$ there is an elementary extension $\mathfrak{N} \succ \mathfrak{M}$ that is recursively saturated and such that $|\mathfrak{N}| = |\mathfrak{M}|$.*

Now to a slightly different notion, that of resplendency. A Σ_1^1 formula is a second-order formula of the form $\exists X\, \varphi(X)$ where X is a set variable and $\varphi(X)$ is a first-order formula in the language extended with the set variable X.

Definition 2.5. $\mathfrak{M}$ is *resplendent* if for all Σ_1^1 $\mathscr{L}_{\mathfrak{M}}$-sentences Φ, such that

$$\mathrm{ElDiag}(\mathfrak{M}) \cup \{\, \Phi \,\}$$

is consistent, we have $\mathfrak{M} \models \Phi$.

In other words $\mathfrak{M}$ is resplendent if as many as possible Σ_1^1 sentences are true and it is recursively saturated if all recursive types are realized. Observe that both these notions apply to all structures, not only models of `PA`. The next theorem follows from work by Kleene [Kle52].

Theorem 2.6. *If $\mathfrak{M}$ is resplendent then it is recursively saturated.*

There is a converse if the model is countable, this result is due to Barwise and Schlipf and independently Ressayre.

Theorem 2.7 ([BS76]). *If $\mathfrak{M}$ is countable and recursively saturated then it is resplendent.*

There are several model theoretic properties which are Σ_1^1, making resplendent models easy to work with. In this thesis we will express the consistency of logics, which are definable in some model, by an Σ_1^1 formula:

$$\exists X \Big(\text{all axioms are in } X \land X \text{ is closed under the inference rules} \\ \land \exists x \big(x \text{ is a formula} \land x \notin X \big) \Big).$$

Hopefully, the reader is now ready to face satisfaction classes.

3

Satisfaction classes

When defining satisfaction classes we have two different approaches to choose from. The historical way (used in [Kra76]) is to look at nonstandard formulas of $\mathscr{L}_A$ and define a satisfaction class to be a set of pairs; the first component being a nonstandard formula of $\mathscr{L}_A$ and the second being a code for a sequence of elements in $\mathfrak{M}$. The intention is that the sequence satisfies the formula, i.e., if we substitute the free variable v_i with the ith element of the sequence then the result is "true" in $\mathfrak{M}$. The disadvantage of this approach is that we need some machinery to handle different ways of getting the "same" formula. We will give an example; let ϵ_0 be $\mathsf{v}_0 \neq \mathsf{v}_0$ and ϵ_{i+1} be $\epsilon_i \vee \epsilon_i$, also let ϵ'_0 be $\mathsf{v}_0 \neq 0$ and ϵ'_{i+1} be $\epsilon'_i \vee \epsilon'_i$, then we clearly want $\langle \epsilon_i, [0] \rangle \in \Sigma$ iff $\langle \epsilon'_i, [0] \rangle \in \Sigma$.

Conjecture 3.1. *There is a satisfaction class Σ, in the sense of [Kra76] and [KKL81], such that $\langle \epsilon_a, [0] \rangle \in \Sigma$ and $\langle \epsilon'_a, [0] \rangle \notin \Sigma$ for some $a \in \mathfrak{M}$.*

We think that by redefining $\mathfrak{M}$-logic to work with pairs of $\mathscr{L}_A$-formulas and elements of $\mathfrak{M}$ it should be possible to prove the conjecture by reproving the results in this chapter.

We are going to define a satisfaction class to be a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences; the intention is that the sentences are those which are "true" in $\mathfrak{M}$. The disadvantage of this approach is that the name 'satisfaction class' seems a bit awkward, a better name would probably be 'truth class', but for historical reasons we will stick with it.

We will use the two notations $x \in \Sigma$ and $\Sigma(x)$ to mean the same thing, i.e., in this case the symbol $\in$ has nothing to do with the coding of finite sets. We hope this will not confuse the reader, but instead make the formulas easier to read.

A comment on the word *class* should be made here. A class in the model theory of arithmetic is a subset C of the domain of the model such that for all a, $C \cap I_{<a}$ is definable (remember that $I_{<a} = \{ x \in \mathfrak{M} \mid x < a \}$). It is just a mere and unfortunate coincidence that the word is used in the term 'satisfaction class.'

Definition 3.2. A satisfaction class Σ is an external subset of $\mathfrak{M}$ satisfying the following conditions in $\mathfrak{M}$:

$$x \in \Sigma \to \mathtt{Sent}(x), \tag{3.1}$$
$$\ulcorner \mathsf{c}_a = \mathsf{c}_b \urcorner \in \Sigma \leftrightarrow a = b, \tag{3.2}$$
$$\ulcorner t = t \urcorner \in \Sigma, \tag{3.3}$$
$$\ulcorner t = r \urcorner \in \Sigma \to \ulcorner r = t \urcorner \in \Sigma, \tag{3.4}$$
$$\ulcorner t = r \urcorner \in \Sigma \land \ulcorner r = s \urcorner \in \Sigma \to \ulcorner t = s \urcorner \in \Sigma, \tag{3.5}$$
$$\ulcorner \mathsf{S}(t) = \mathsf{c}_a \urcorner \in \Sigma \leftrightarrow \exists x \big(\ulcorner t = \mathsf{c}_x \urcorner \in \Sigma \land \mathsf{S}(x) = a \big), \tag{3.6}$$
$$\ulcorner t + r = \mathsf{c}_a \urcorner \in \Sigma \leftrightarrow \exists x, y \big(\ulcorner t = \mathsf{c}_x \urcorner \in \Sigma \land \ulcorner r = \mathsf{c}_y \urcorner \in \Sigma \land x + y = a \big), \tag{3.7}$$
$$\ulcorner t \cdot r = \mathsf{c}_a \urcorner \in \Sigma \leftrightarrow \exists x, y \big(\ulcorner t = \mathsf{c}_x \urcorner \in \Sigma \land \ulcorner r = \mathsf{c}_y \urcorner \in \Sigma \land x \cdot y = a \big), \tag{3.8}$$
$$\ulcorner t = r \urcorner \in \Sigma \leftrightarrow \exists x \big(\ulcorner t = \mathsf{c}_x \urcorner \in \Sigma \land \ulcorner r = \mathsf{c}_x \urcorner \in \Sigma \big), \tag{3.9}$$
$$\ulcorner \neg \varphi \urcorner \in \Sigma \leftrightarrow \varphi \notin \Sigma, \tag{3.10}$$
$$\ulcorner \varphi \lor \psi \urcorner \in \Sigma \leftrightarrow \big(\varphi \in \Sigma \lor \psi \in \Sigma \big) \quad \text{and} \tag{3.11}$$
$$\ulcorner \exists \mathsf{v}_i \, \gamma \urcorner \in \Sigma \leftrightarrow \exists x \big(\ulcorner \gamma[\mathsf{c}_x / \mathsf{v}_i] \urcorner \in \Sigma \big), \tag{3.12}$$

for all closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms t, r and s and all ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences φ, ψ and $\exists \mathsf{v}_i \gamma$.

The clauses (3.2)-(3.9) take care of the atomic formulas and we could replace them with the single clause

$$t = r \in \Sigma \leftrightarrow \mathfrak{M} \models \mathtt{val}(t) = \mathtt{val}(r),$$

but that would yield a stronger notion. We will look more closely at this later in Chapter 5.

Observe that there is a first-order formula $\mathtt{SatCl}(X)$ in the language $\mathscr{L}_A \cup \{ X \}$ such that for any subset A of $\mathfrak{M}$ A is a satisfaction class iff $\mathfrak{M} \models \mathtt{SatCl}(A)$. The formula $\mathtt{SatCl}(X)$ is the conjunction of the universal closures (with the universal quantifiers restricted to closed terms and sentences) of (3.1)-(3.12), which all are first-order properties.

Proposition 3.3. *Let Σ be a satisfaction class, then the following is true in the structure $\langle \mathfrak{M}, \Sigma \rangle$:*

$$\begin{aligned}
\ulcorner \varphi \land \psi \urcorner \in \Sigma &\leftrightarrow \varphi \in \Sigma \land \psi \in \Sigma, \\
\ulcorner \varphi \to \psi \urcorner \in \Sigma &\leftrightarrow \big(\varphi \in \Sigma \to \psi \in \Sigma \big), \\
\ulcorner \varphi \leftrightarrow \psi \urcorner \in \Sigma &\leftrightarrow \big(\varphi \in \Sigma \leftrightarrow \psi \in \Sigma \big), \quad \text{and} \\
\ulcorner \forall \mathsf{v}_i \, \gamma \urcorner \in \Sigma &\leftrightarrow \forall x \big(\ulcorner \gamma[\mathsf{c}_x / \mathsf{v}_i] \urcorner \in \Sigma \big),
\end{aligned}$$

for all ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences φ, ψ and $\forall \mathsf{v}_i \, \gamma$.

Proposition 3.4. *If Σ is a satisfaction class, φ an $\mathscr{L}_{\mathfrak{M}}$-sentence and $\mathfrak{M} \models \varphi$ then $\varphi \in \Sigma$, i.e.,* $\mathrm{ElDiag}(\mathfrak{M}) \subseteq \Sigma$.

Proof. We first prove the statement for sentences φ of the form $t = \mathsf{c}_a$ for some closed term t and some $a \in \mathfrak{M}$, by induction on the construction of t.

If t is c_b then $\mathfrak{M} \vDash \mathsf{c}_b = \mathsf{c}_a$ so $\mathsf{c}_b = \mathsf{c}_a \in \Sigma$ by (3.2).

Suppose t is $\mathsf{S}(t')$ for some closed term t'; then $\mathfrak{M} \vDash t' = \mathsf{c}_b$ for some $b \in \mathfrak{M}$ and by the induction hypothesis we have $t' = \mathsf{c}_b \in \Sigma$. Also $\mathsf{S}^{\mathfrak{M}}(b) = a$ so $\mathsf{S}(t') = \mathsf{c}_a \in \Sigma$ by (3.6).

Suppose t is $t_1 + t_2$, then there are b, $b' \in \mathfrak{M}$ such that $\mathfrak{M} \vDash t_1 = \mathsf{c}_b$ and $\mathfrak{M} \vDash t_2 = \mathsf{c}_{b'}$. By the induction hypothesis we have $t_1 = \mathsf{c}_b \in \Sigma$ and $t_2 = \mathsf{c}_{b'} \in \Sigma$. We know that $b +^{\mathfrak{M}} b' = a$, so $t_1 + t_2 = \mathsf{c}_a \in \Sigma$ by (3.7). The case when t is $t_1 \cdot t_2$ is treated in a similar way.

Suppose now $\mathfrak{M} \vDash t_1 = t_2$ then there is a $b \in \mathfrak{M}$ such that $\mathfrak{M} \vDash t_1 = b$ and $\mathfrak{M} \vDash t_2 = b$, so $t_1 = \mathsf{c}_b \in \Sigma$ and $t_2 = \mathsf{c}_b \in \Sigma$ which by (3.4) and (3.5) implies $t_1 = t_2 \in \Sigma$.

If $\mathfrak{M} \vDash t_1 \neq t_2$ then there are b_1, $b_2 \in \mathfrak{M}$ such that

$$\mathfrak{M} \vDash t_1 = b_1 \wedge t_2 = b_2 \wedge b_1 \neq b_2,$$

so $t_1 = \mathsf{c}_{b_1} \in \Sigma$ and $t_2 = \mathsf{c}_{b_2} \in \Sigma$. By (3.2) we also have $\mathsf{c}_{b_1} \neq \mathsf{c}_{b_2} \in \Sigma$ and by (3.4) and (3.5) we have $t_1 \neq t_2 \in \Sigma$.

We have proved that $\mathrm{Diag}(\mathfrak{M}) \subseteq \Sigma$. We should prove that if φ is an $\mathscr{L}_{\mathfrak{M}}$-sentence then

$$\mathfrak{M} \vDash \varphi \quad \Leftrightarrow \quad \varphi \in \Sigma$$

by induction on φ. The case when φ is atomic is proven. The induction step is easy and left to the reader. ☐

Remark 3.5. By Tarski's theorem of the undefinability of truth a satisfaction class Σ is not definable. Moreover $\Sigma \cap I_{<a}$ for $a > \omega$ is not definable, since this would also yield a truth definition, therefore Σ is not a class.

Given a set of ${}^*\!\mathscr{L}_{\mathfrak{M}}$-sentences X, define the binary relation $\sim_X$ (or just $\sim$ if X is understood from the context) on the set of closed ${}^*\!\mathscr{L}_{\mathfrak{M}}$-terms, $\texttt{ClTerm}(\mathfrak{M})$, as follows

$$t_1 \sim_X t_2 \quad \text{iff} \quad t_1 = t_2 \in X.$$

If $\sim_X$ is an equivalence relation let $\texttt{ClTerm}(\mathfrak{M})/\!\sim_X$ be the set of equivalence classes $\bar{t}$ of $\texttt{ClTerm}(\mathfrak{M})$ and define the functions $\mathsf{S}^{\mathfrak{M}_X}$, $+^{\mathfrak{M}_X}$ and $\cdot^{\mathfrak{M}_X}$ by

$$\begin{aligned} \mathsf{S}^{\mathfrak{M}_X}(\,\bar{t}\,) &=_{\mathrm{df}} \overline{\mathsf{S}(t)}, \\ \bar{t} +^{\mathfrak{M}_X} \bar{r} &=_{\mathrm{df}} \overline{t + r} \quad \text{and} \\ \bar{t} \cdot^{\mathfrak{M}_X} \bar{r} &=_{\mathrm{df}} \overline{t \cdot r} \end{aligned}$$

if they all are well-defined. This defines an $\mathscr{L}_A$-structure

$$(\texttt{ClTerm}(\mathfrak{M})/\!\sim_X, \mathsf{S}^{\mathfrak{M}_X}, +^{\mathfrak{M}_X}, \cdot^{\mathfrak{M}_X}, \bar{0}).$$

Let $\mathfrak{M}_X$ denote this structure, when it is well-defined.

The following equivalent definition of a satisfaction class may be well worth notice; it will play the main role of Chapter 4.

Proposition 3.6. *Σ is a satisfaction class iff it is a unary predicate on $\mathfrak{M}$ such that $\sim_\Sigma$ is an equivalence relation, $\mathfrak{M}_\Sigma$ is well-defined, the canonical map*

$$f : \mathfrak{M} \to \mathfrak{M}_\Sigma, \quad a \mapsto \overline{\mathsf{c}_a}$$

is an isomorphism and

$$x \in \Sigma \to \mathtt{Sent}(x), \tag{3.1}$$

$$\ulcorner\neg\varphi\urcorner \in \Sigma \leftrightarrow \varphi \notin \Sigma, \tag{3.10}$$

$$\ulcorner\varphi \vee \psi\urcorner \in \Sigma \leftrightarrow \big(\varphi \in \Sigma \vee \psi \in \Sigma\big), \quad \text{and} \tag{3.11}$$

$$\ulcorner\exists \mathsf{v}_i\, \gamma\urcorner \in \Sigma \leftrightarrow \exists x\big(\ulcorner\gamma[\mathsf{c}_x/\mathsf{v}_i]\urcorner \in \Sigma\big) \tag{3.12}$$

holds for all ${}^*\mathscr{L}_\mathfrak{M}$-sentences φ, ψ and $\exists \mathsf{v}_i\, \gamma$.

Proof. Assume Σ is a satisfaction class. By (3.3)-(3.5) it is clear that $\sim_\Sigma$ is an equivalence relation and by (3.6)-(3.8) $\mathfrak{M}_\Sigma$ is well-defined. It is also clear that the mapping f is a bijection since if $a \neq b$ then $\mathsf{c}_a = \mathsf{c}_b \notin \Sigma$ so $f(a) \neq f(b)$ and if $t \in \mathtt{ClTerm}(\mathfrak{M})$ then $t = t \in \Sigma$, so by (3.9) there exists $m \in \mathfrak{M}$ such that $t = \mathsf{c}_m \in \Sigma$, i.e., $f(m) = \overline{t}$. We also have

$$f(\mathsf{S}^\mathfrak{M}(a)) = \overline{\mathsf{c}_{\mathsf{S}^\mathfrak{M}(a)}} = \overline{\mathsf{S}(\mathsf{c}_a)} = \mathsf{S}^{\mathfrak{M}_X}(\overline{\mathsf{c}_a}) = \mathsf{S}^{\mathfrak{M}_X}(f(a)),$$
$$f(a +^\mathfrak{M} b) = \overline{\mathsf{c}_{a+^\mathfrak{M} b}} = \overline{\mathsf{c}_a + \mathsf{c}_b} = \overline{\mathsf{c}_a} +^{\mathfrak{M}_X} \overline{\mathsf{c}_b} = f(a) +^{\mathfrak{M}_X} f(b) \quad \text{and}$$
$$f(a \cdot^\mathfrak{M} b) = \overline{\mathsf{c}_{a\cdot^\mathfrak{M} b}} = \overline{\mathsf{c}_a \cdot \mathsf{c}_b} = \overline{\mathsf{c}_a} \cdot^{\mathfrak{M}_X} \overline{\mathsf{c}_b} = f(a) \cdot^{\mathfrak{M}_X} f(b),$$

so f is really an isomorphism.

Assume now that Σ satisfies the conditions above.

We have to prove conditions (3.2) and (3.6)-(3.9) in the definition of satisfaction classes. The other follows immediately.

(3.2) It follows from the injectivity of f, since if $a \neq b \in \mathfrak{M}$ then $f(a) \neq f(b)$ so $\overline{\mathsf{c}_a} \neq \overline{\mathsf{c}_b}$, i.e., $\mathsf{c}_a \neq \mathsf{c}_b \in \Sigma$.

(3.6) Assume $\mathsf{S}(t) = \mathsf{c}_a \in \Sigma$, we have to find m such that $t = \mathsf{c}_m \in \Sigma$ and $\mathsf{S}^\mathfrak{M}(m) = a$. We have

$$a = f^{-1}(\overline{\mathsf{c}_a}) = f^{-1}(\overline{\mathsf{S}(t)}) = f^{-1}(\mathsf{S}^{\mathfrak{M}_X}(\,\overline{t}\,)) = \mathsf{S}^\mathfrak{M}(f^{-1}(\,\overline{t}\,)),$$

so $m = f^{-1}(\,\overline{t}\,)$ works since $\overline{t} = f(m)$ implies $t = \mathsf{c}_m \in \Sigma$. On the other hand assume that there exists m such that $t = \mathsf{c}_m \in \Sigma$ and $\mathsf{S}^\mathfrak{M}(m) = a$ then

$$\overline{\mathsf{S}(t)} = \mathsf{S}^{\mathfrak{M}_X}(\,\overline{t}\,) = \mathsf{S}^{\mathfrak{M}_X}(\overline{\mathsf{c}_m}) = \mathsf{S}^{\mathfrak{M}_X}(f(m)) = f(\mathsf{S}^\mathfrak{M}(m)) = \overline{\mathsf{c}_a}.$$

(3.7) Assume $t + r = \mathsf{c}_a \in \Sigma$ and let $m = f^{-1}(\,\overline{t}\,)$ and $m' = f^{-1}(\overline{r})$ then $a = m +^\mathfrak{M} m'$ and $t = \mathsf{c}_m \in \Sigma$ and $r = \mathsf{c}_{m'} \in \Sigma$. If $m +^\mathfrak{M} m' = a$, $t = \mathsf{c}_m \in \Sigma$ and $r = \mathsf{c}_{m'} \in \Sigma$ then $f^{-1}(\,\overline{t}\,) = m$ and $f^{-1}(\overline{r}) = m'$ so

$$a = f^{-1}(\overline{t} +^{\mathfrak{M}_X} \overline{r}) = f^{-1}(\overline{t + r}).$$

In other words $f(a) = \overline{t + r}$, i.e., $t + r = \mathsf{c}_a \in \Sigma$. (3.8) Similar as above.

(3.9) Assume $t = r \in \Sigma$ then $f^{-1}(\,\overline{t}\,) = f^{-1}(\overline{r}) = m$ so $t = \mathsf{c}_m$, $r = \mathsf{c}_m \in \Sigma$ and if $t = \mathsf{c}_m$, $r = \mathsf{c}_m \in \Sigma$ then $\overline{t} = \overline{\mathsf{c}_m} = \overline{r}$, so $t = r \in \Sigma$. ∎

Question 3.7. *Are satisfaction classes built up by two "parts;" one with closed terms and equality and one with "the rest," in the following sense: Given a relation* $\sim$ *satisfying the conditions in Proposition 3.6 is there a satisfaction class* Σ *such that the relation* $\sim_\Sigma$ *coincides with* $\sim$*?*

A partial answer to this question is given in Proposition 3.58.

3.1 Inductive partial satisfaction classes

This section is included as an introduction to satisfaction classes; the results will not be used later and may therefore be skipped. For simplicity the satisfaction classes in this section will all include the set

$$\{\, t = r \mid \mathfrak{M} \models \mathtt{val}(t) = \mathtt{val}(r) \,\}.$$

Therefore, in this section, $\Sigma \subseteq \mathfrak{M}$ is a satisfaction class iff

$$\langle \mathfrak{M}, \Sigma \rangle \models \forall x \big(x \in \Sigma \to \mathtt{Sent}(x)\big) \wedge \forall x \Psi(\Sigma, x),$$

where $\Psi(\Sigma, x)$ is the formula

$$\begin{aligned} &\forall y, z \big[\mathtt{ClTerm}(y) \wedge \mathtt{ClTerm}(z) \wedge x = \ulcorner y = z \urcorner \to \big(x \in \Sigma \leftrightarrow \mathtt{val}(y) = \mathtt{val}(z)\big)\big] \\ &\quad \wedge \forall y \big[\mathtt{Sent}(y) \wedge x = \ulcorner \neg y \urcorner \to \big(x \in \Sigma \leftrightarrow y \notin \Sigma\big)\big] \\ &\quad \wedge \forall y, z \big[\mathtt{Sent}(y) \wedge \mathtt{Sent}(z) \wedge x = \ulcorner y \vee z \urcorner \to \big(x \in \Sigma \leftrightarrow y \in \Sigma \vee z \in \Sigma\big)\big] \\ &\quad \wedge \forall y, i \big[\mathtt{Sent}(\exists \mathsf{v}_i\, y) \wedge x = \ulcorner \exists \mathsf{v}_i\, y \urcorner \to \big(x \in \Sigma \leftrightarrow \exists z \ulcorner y[\mathsf{c}_z/\mathsf{v}_i] \urcorner \in \Sigma\big)\big]. \end{aligned}$$

We will assume that if ψ is a subformula or subterm of φ then $\ulcorner \psi \urcorner \leqslant \ulcorner \varphi \urcorner$; under the assumption that the Gödel number of a formula is the sequence of Gödel numbers of the symbols in the formula this is true.

Definition 3.8. A set $\Sigma \subseteq \mathfrak{M}$ is a *partial satisfaction class* if there exists $c \in \mathfrak{M} \setminus \omega$ such that

$$\langle \mathfrak{M}, \Sigma \rangle \models \forall x \big(x \in \Sigma \to \mathtt{Sent}(x)\big) \wedge \forall \varphi {<} c\, \forall x\, \Psi(\Sigma, \ulcorner \varphi[x/\mathsf{v}] \urcorner).^1$$

A partial satisfaction class is *inductive* if full induction in the language $\mathscr{L}_A \cup \{\, \Sigma \,\}$ holds, i.e., if

$$\langle \mathfrak{M}, \Sigma \rangle \models \varphi(0) \wedge \forall x \big(\varphi(x) \to \varphi(\mathsf{S}(x))\big) \to \forall x\, \varphi(x)$$

for every $\mathscr{L}_{\mathfrak{M}} \cup \{\, \Sigma \,\}$-formula $\varphi(x)$.

Proposition 3.9. *If* φ *is an* $\mathscr{L}_{\mathfrak{M}}$*-sentence and* Σ *a partial satisfaction class then*

$$\varphi \in \Sigma \quad \textit{iff} \quad \mathfrak{M} \models \varphi.$$

[1] See Definition 3.31 for the definition of $\varphi[a/\mathsf{v}]$.

Proof. For atomic sentences φ the proposition follows trivial from the definition of partial satisfaction classes. The rest is an easy induction on the construction of φ. $\square$

Theorem 3.10. *Every countable recursively saturated model $\mathfrak{M}$ of* PA *admits an inductive partial satisfaction class.*

Proof. Since $\mathfrak{M}$ is countable and recursively saturated it is resplendent. Therefore we only need to show that the theory

$$\begin{aligned}&\mathrm{ElDiag}(\mathfrak{M})\\ &\quad+\{\,\varphi(0)\wedge\forall x\big(\varphi(x)\to\varphi(\mathsf{S}(x))\big)\to\forall x\,\varphi(x)\,\big|\,\varphi(x)\text{ an }\mathscr{L}_{\mathfrak{M}}\cup\{\,\Sigma\,\}\text{-formula}\,\}\\ &\qquad\qquad+\forall x\big(x\in\Sigma\to\mathtt{Sent}(x)\big)+\{\,\forall x\,\Psi(\Sigma,\varphi[x/\mathsf{v}])\,|\,\varphi\text{ an }\mathscr{L}_A\text{-formula}\,\}\end{aligned}$$

is consistent. If this theory is consistent then clearly any Σ satisfying it will be an inductive partial satisfaction class by a simple overspill argument since

$$\forall n{\in}\omega\ \langle\mathfrak{M},\Sigma\rangle\vDash\forall x{<}n\,\forall x\,\Psi(\Sigma,\varphi[x/\mathsf{v}]).$$

To prove the consistency take a finite subset of the theory, it will at most involve a finite number of standard formulas φ in the scheme

$$\forall x\,\Psi(\Sigma,\varphi[x/\mathsf{v}]).$$

Since it is a finite number they will all be Σ_k for some $k\in\omega$, therefore we can define

$$\Sigma=_{\mathrm{df}}\ \{\,a\in\mathfrak{M}\,|\,\mathfrak{M}\vDash\mathtt{Tr}_{\Sigma_k}(a)\,\}\,.$$ $\square$

Theorem 3.11. *If $\mathfrak{M}\vDash$* PA *is a nonstandard model admitting an inductive partial satisfaction class Σ then $\mathfrak{M}$ is recursively saturated.*

Proof. Suppose $p(x)$ is a recursive type in $\mathfrak{M}$ and $\bar{a}\in\mathfrak{M}$ are the parameters of the type. The type, being recursive, is coded by some $b\in\mathfrak{M}$. Since

$$\forall k{\in}\omega\ \langle\mathfrak{M},\Sigma\rangle\vDash\exists x\forall\varphi{<}k\big(\varphi\in b\to\ulcorner\varphi\big[[x,\bar{a}]/\mathsf{v}\big]\urcorner\in\Sigma\big)$$

we can use overspill (it is here we are using that Σ is inductive) and get

$$\langle\mathfrak{M},\Sigma\rangle\vDash\exists x\forall\varphi{<}c\big(\varphi\in b\to\ulcorner\varphi\big[[x,\bar{a}]/\mathsf{v}\big]\urcorner\in\Sigma\big)$$

for some $c\in\mathfrak{M}\setminus\omega$. This gives us an $x\in\mathfrak{M}$ realizing $p(x)$. $\square$

Remark 3.12. We are going to do the same sort of argument when Σ does not satisfy induction. The argument there involves more work.

Combining these two theorems we get the following.

Theorem 3.13. *A nonstandard countable model of* PA *is recursively saturated iff it admits an inductive partial satisfaction class.*

3.2 Construction of satisfaction classes

We will prove the following theorem.

Theorem 3.14 ([KKL81]). *If $\mathfrak{M}$ is countable and recursively saturated then $\mathfrak{M}$ admits a satisfaction class.*

To prove it we will define a logic, called $\mathfrak{M}$-logic, which we will prove to be consistent (if $\mathfrak{M}$ is recursively saturated). Then we construct a maximally consistent set of sentences in this logic, which turns out to be a satisfaction class. The idea to use $\mathfrak{M}$-logic is due to Jeff Paris.

3.2.1 $\mathfrak{M}$-logic

As formulas in this logic we will consider all ${}^*\mathscr{L}_{\mathfrak{M}}$*-sentences*. We will consider a formal deduction system that derives actually finite sets of formulas, usually denoted by upper case Greek letters. The intention is that the set Γ should be read as $\bigvee\Gamma$, i.e., the disjunction of the sentences in Γ. The symbol φ will denote the singleton set $\{\,\varphi\,\}$ and Γ, φ will denote $\Gamma \cup \{\,\varphi\,\}$. The axioms for the deductive system are:

$$\varphi, \neg\varphi \tag{Axiom1}$$
$$\mathsf{c}_a \neq \mathsf{c}_b \quad \text{if } a \neq b \tag{Axiom2}$$
$$t = t \tag{Axiom3}$$
$$t \neq r, r = t \tag{Axiom4}$$
$$t \neq r, r \neq s, t = s \tag{Axiom5}$$
$$t \neq r, \mathsf{S}(t) = \mathsf{S}(r) \tag{Axiom6}$$
$$t \neq t', r \neq r', t + r = t' + r' \tag{Axiom7}$$
$$t \neq t', r \neq r', t \cdot r = t' \cdot r' \tag{Axiom8}$$
$$\mathsf{S}(\mathsf{c}_a) = \mathsf{c}_{\mathsf{S}^{\mathfrak{M}}(a)} \tag{Axiom9}$$
$$\mathsf{c}_a + \mathsf{c}_b = \mathsf{c}_{a +^{\mathfrak{M}} b} \tag{Axiom10}$$
$$\mathsf{c}_a \cdot \mathsf{c}_b = \mathsf{c}_{a \cdot^{\mathfrak{M}} b} \tag{Axiom11}$$
$$\exists \mathsf{v}_0 (t = \mathsf{v}_0) \tag{Axiom12}$$

where φ is an arbitrary ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence and t, r and s are arbitrary closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms. The inference rules are the following:

$$\frac{\Gamma}{\Gamma, \varphi} \tag{Weak}$$

$$\frac{\Gamma, \varphi}{\Gamma, \varphi \vee \psi} \tag{$\vee$I1}$$

$$\frac{\Gamma, \psi}{\Gamma, \varphi \vee \psi} \tag{$\vee$I2}$$

$$\frac{\Gamma, \neg\varphi \quad \Gamma, \neg\psi}{\Gamma, \neg(\varphi \vee \psi)} \tag{$\vee$I3}$$

$$\frac{\Gamma, \varphi}{\Gamma, \neg\neg\varphi} \tag{$\neg$I}$$

$$\frac{\Gamma, \varphi \quad \Gamma, \neg\varphi}{\Gamma} \tag{Cut}$$

$$\frac{\Gamma, \varphi[\mathsf{c}_a/\mathsf{v}_i]}{\Gamma, \exists \mathsf{v}_i\, \varphi} \tag{$\exists$I}$$

$$\frac{\ldots\ \Gamma, \neg\varphi[\mathsf{c}_a/\mathsf{v}_i]\ \ldots\ {}_{a\in\mathfrak{M}}}{\Gamma, \neg\exists \mathsf{v}_i\, \varphi} \tag{$\mathfrak{M}$-rule}$$

where Γ is an arbitrary finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, φ is an arbitrary ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence and the premises in $\mathfrak{M}$-rule means $\Gamma, \neg\varphi[\mathsf{c}_a/\mathsf{v}_i]$ for all $a \in \mathfrak{M}$.

It might be worth noticing that it is only Axiom2, Axiom9, Axiom10, Axiom11 and $\mathfrak{M}$-rule that depends on the model $\mathfrak{M}$.

At a first glance Axiom12 seems to be an ugly duckling, but it turns out to be of great importance. We will discuss its importance in Chapter 4.

Observe the connection with ω-logic: the logic we get by replacing $\mathfrak{M}$-rule with

$$\frac{\ldots\ \Gamma, \varphi[\mathsf{c}_a/\mathsf{v}_i]\ \ldots\ {}_{a\in\mathfrak{M}}}{\Gamma, \forall \mathsf{v}_i\, \varphi}$$

is essentially equivalent to $\mathfrak{M}$-logic.

The definition of a proof follows the usual one. A proof is a tree where the nodes are finite sets of sentences, the leaves are axioms and the root is the conclusion and each edge from one node to another follows one of the inference rules. For example;

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\text{Axiom2}}{\varphi, \neg\varphi}}{\varphi, \psi, \neg\varphi}\text{Weak} \quad \dfrac{\dfrac{\text{Axiom2}}{\psi, \neg\psi}}{\varphi, \psi, \neg\psi}\text{Weak}}{\neg(\varphi \vee \psi), \varphi, \psi}\vee\text{I3}}{\neg(\varphi \vee \psi), \varphi, \psi \vee \varphi}\vee\text{I2}}{\neg(\varphi \vee \psi), \psi \vee \varphi}\vee\text{I1}}{\dfrac{\neg(\varphi \vee \psi) \vee (\psi \vee \varphi), \psi \vee \varphi}{\neg(\varphi \vee \psi) \vee (\psi \vee \varphi)}\vee\text{I1}}\vee\text{I2}$$

is a proof of the commutativity of $\vee$, i.e., of $(\varphi \vee \psi) \to (\psi \vee \varphi)$. Figure 3.1 shows the corresponding tree.

The notation $\vdash^p_{\mathfrak{M}} \Gamma$ will mean that p is a proof in $\mathfrak{M}$-logic with conclusion Γ and $\vdash_{\mathfrak{M}} \Gamma$ means that there is a proof p such that $\vdash^p_{\mathfrak{M}} \Gamma$. Please observe that the proofs are external objects and might *not* be definable in $\mathfrak{M}$, they are either finite or infinite trees.

If Λ is a (external, finite or infinite) set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences then $\Lambda \vdash^p_{\mathfrak{M}} \Gamma$ means that p is a proof of Γ in $\mathfrak{M}$-logic with the added axioms

$$\varphi \qquad \text{if } \varphi \in \Lambda. \tag{AxiomΛ}$$

$$
\begin{tikzcd}
\varphi, \neg\varphi \arrow[d, dash, "\text{Weak}"] & & \psi, \neg\psi \arrow[d, dash, "\text{Weak}"] \\
\varphi, \psi, \neg\varphi \arrow[dr, dash, "\vee\text{I3}"] & & \varphi, \psi, \neg\psi \arrow[dl, dash, "\vee\text{I3}"'] \\
& \neg(\varphi \vee \psi), \varphi, \psi \arrow[d, dash, "\vee\text{I2}"] & \\
& \neg(\varphi \vee \psi), \varphi, \psi \vee \varphi \arrow[d, dash, "\vee\text{I1}"] & \\
& \neg(\varphi \vee \psi), \psi \vee \varphi \arrow[d, dash, "\vee\text{I2}"] & \\
& \neg(\varphi \vee \psi) \vee (\psi \vee \varphi), \psi \vee \varphi \arrow[d, dash, "\vee\text{I1}"] & \\
& \neg(\varphi \vee \psi) \vee (\psi \vee \varphi) &
\end{tikzcd}
$$

Figure 3.1: *The tree form of the proof of* $(\varphi \vee \psi) \to (\psi \vee \varphi)$.

$\Lambda \vdash_{\mathfrak{M}} \Gamma$ means that there is a proof p such that $\Lambda \vdash^{p}_{\mathfrak{M}} \Gamma$.

If $\Lambda \vdash^{p}_{\mathfrak{M}} \Gamma$ then $|p|$ denotes the height of p which is the height of p looked on as a tree, or in other words;

Definition 3.15. The *height*, $|p|$, of a proof p is defined as follows; if p is an axiom then $|p| = 0$, if the last inference rule in p is of the form

$$\frac{\begin{array}{ccccc} p_1 & p_2 & & p_i & \\ \Gamma_1 & \Gamma_2 & \dots & \Gamma_i & \dots \end{array}}{\Gamma}$$

(there are either one, two or infinitely many premises) then

$$|p| = \sup_i (|p_i| + 1).$$

The height $|p|$ is an ordinal number in the "real world." By the $\mathfrak{M}$-rule $|p|$ might be infinite. For example Figure 3.2 shows a proof of height ω. As we will see later, the question of whether there are sets Γ provable in $\mathfrak{M}$-logic but not provable by a proof of finite height is equivalent to the question if $\mathfrak{M}$-logic is consistent.

If α is an ordinal number then $\Lambda \vdash^{\alpha}_{\mathfrak{M}} \Gamma$ means that there is a proof p such that $|p| < \alpha$ and $\Lambda \vdash^{p}_{\mathfrak{M}} \Gamma$, e.g., $\vdash^{\omega}_{\mathfrak{M}} \Gamma$ means that there is a finite height proof of Γ.

As an alternative we may define

$$\vdash^{\alpha}_{\mathfrak{M}} \Gamma \quad \text{iff} \quad \Gamma \in I_\alpha$$

where I_1 is the set of axioms of $\mathfrak{M}$-logic,

$$I_{\alpha+1} = \{\, \Gamma \mid \Gamma \text{ is the result of applying one of the inference rules to sets in } I_\alpha \,\}$$

and

$$I_\lambda = \bigcup_{\alpha<\lambda} I_\alpha$$

for limit ordinals λ. Then we could have defined $\vdash_{\mathfrak{M}} \Gamma$ to hold if there exists an ordinal α such that $\vdash^{\alpha}_{\mathfrak{M}} \Gamma$. This alternative definition avoids mentioning trees, simplifying things.

Lemma 3.16. *Let Λ be an arbitrary set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences and Δ and Γ be finite sets of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences. Then*

1. $\Lambda, \Delta \vdash_{\mathfrak{M}} \Gamma$ *iff* $\Lambda \vdash_{\mathfrak{M}} \neg\Delta, \Gamma$ *where* $\neg\Delta = \{\, \neg\varphi \mid \varphi \in \Delta \,\}$.
2. $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi \vee \psi$ *iff* $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi$.
3. $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi$ *iff* $\Lambda \vdash_{\mathfrak{M}} \Gamma, \neg\neg\varphi$.

Proof. 1. If $\Delta = \emptyset$, the lemma is trivial. Suppose $\Delta = \{\varphi\}$, then if $\Lambda \vdash_{\mathfrak{M}} \neg\varphi, \Gamma$ it is clear that $\Lambda, \varphi \vdash_{\mathfrak{M}} \neg\varphi, \Gamma$ and $\Lambda, \varphi \vdash_{\mathfrak{M}} \varphi, \Gamma$ (by AxiomΛ and Weak) so by Cut we get $\Lambda, \varphi \vdash_{\mathfrak{M}} \Gamma$.

$$
\dfrac{\dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_0}=\mathsf{c}_{a_0}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_1}=\mathsf{c}_{a_1}} \quad \dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_1}=\mathsf{c}_{a_1}}}{\mathsf{c}_{a_1}=\mathsf{c}_{a_1},\neg\varphi}}{\mathsf{c}_{a_1}=\mathsf{c}_{a_1},\varphi} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2}} \quad \dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2}}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2},\neg\varphi}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2},\varphi} \quad \dfrac{\dfrac{\text{Axiom3}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2}}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2},\neg\varphi}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2}} \qquad \dots}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2},\varphi}}{\mathsf{c}_{a_2}=\mathsf{c}_{a_2}}}{\neg\mathsf{c}_{a_2}\neq\mathsf{c}_{a_2}}}{\mathsf{c}_{a_1}=\mathsf{c}_{a_1}}}{\neg\mathsf{c}_{a_1}\neq\mathsf{c}_{a_1}}}{\neg\mathsf{c}_{a_0}\neq\mathsf{c}_{a_0}}}{\neg\exists\mathsf{v}_1(\mathsf{v}_1\neq\mathsf{v}_1)}
$$

Figure 3.2: *An example of a proof of height ω. Here $\{a_i\}_{i\in\omega}$ is an enumeration of the domain of $\mathfrak{M}$ and φ is, for example,* $\mathsf{c}_{a_0}=\mathsf{c}_{a_0}$.

On the other hand assume that

$$\Lambda, \varphi \vdash^{p}_{\mathfrak{M}} \Gamma.$$

We will do induction on $|p|$; if $|p| = 0$ then Γ is an axiom, if $\Gamma \neq \{\, \varphi \,\}$ it is clear that $\Lambda \vdash_{\mathfrak{M}} \Gamma, \neg\varphi$ by Weak and if $\Gamma = \{\, \varphi \,\}$ then $\Gamma, \neg\varphi$ is Axiom1.

If $|p| \geqslant 1$ then the last inference in p is of the form

$$\frac{\Gamma_0 \quad \Gamma_1 \quad \ldots}{\Gamma}$$

and by the induction hypothesis we know that $\Lambda \vdash_{\mathfrak{M}} \Gamma_i, \neg\varphi$, so by the inference

$$\frac{\Gamma_0, \neg\varphi \quad \Gamma_1, \neg\varphi \quad \ldots}{\Gamma, \neg\varphi}$$

we get $\Lambda \vdash_{\mathfrak{M}} \Gamma, \neg\varphi$.

To prove the statement in the general case when $|\Delta| > 1$ we iterate, moving one formula at a time by induction on $|\Delta|$.

2. Suppose $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi$, then by $\vee$I1 $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi \vee \psi, \psi$ and by $\vee$I2 $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi \vee \psi$.

If $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi \vee \psi$ then by Axiom1

$$\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi, \neg\varphi \quad \text{and} \quad \Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi, \neg\psi,$$

so by $\vee$3-rule we have $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi, \neg(\varphi \vee \psi)$ and then by Weak and Cut we get $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \psi$.

3. One direction is $\neg$I. Suppose $\Lambda \vdash_{\mathfrak{M}} \Gamma, \neg\neg\varphi$, then by Weak we have

$$\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \neg\neg\varphi$$

and by using Axiom1 and Weak we get

$$\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi, \neg\varphi,$$

so, finally, Cut gives us $\Lambda \vdash_{\mathfrak{M}} \Gamma, \varphi$. $\square$

Definition 3.17. If Γ is a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, then it is said to be *consistent in* $\mathfrak{M}$*-logic* if $\Gamma \nvdash_{\mathfrak{M}} \emptyset$.

Proposition 3.18. *If Γ is a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences then the following statements are equivalent:*

1. *Γ is consistent in $\mathfrak{M}$-logic.*
2. *There exists a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence φ such that $\Gamma \nvdash_{\mathfrak{M}} \varphi$.*
3. *For all ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences φ either $\Gamma \nvdash_{\mathfrak{M}} \varphi$ or $\Gamma \nvdash_{\mathfrak{M}} \neg\varphi$.*

Proof. 1 $\Rightarrow$ 2. Suppose Γ proves all ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences. Let φ be any such, then $\Gamma \vdash_{\mathfrak{M}} \varphi$ and $\Gamma \vdash_{\mathfrak{M}} \neg\varphi$. By Cut $\Gamma \vdash_{\mathfrak{M}} \emptyset$, i.e., Γ is inconsistent.

2 $\Rightarrow$ 3. Suppose there is a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence φ such that $\Gamma \vdash_{\mathfrak{M}} \varphi$ and $\Gamma \vdash_{\mathfrak{M}} \neg\varphi$ then by Cut $\Gamma \vdash_{\mathfrak{M}} \emptyset$ and so by Weak $\Gamma \vdash_{\mathfrak{M}} \psi$ for any ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence ψ.

3 $\Rightarrow$ 1. Suppose Γ is inconsistent; by Weak Γ proves every ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence, so there certainly exists a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence φ such that $\Gamma \vdash_{\mathfrak{M}} \varphi$ and $\Gamma \vdash_{\mathfrak{M}} \neg\varphi$. $\square$

It follows from Lemma 3.16 that if Γ is consistent (in $\mathfrak{M}$-logic) and φ is a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence then Γ, φ or $\Gamma, \neg\varphi$ is consistent (in $\mathfrak{M}$-logic).

Proposition 3.19. *If* $\varphi \in \mathrm{ElDiag}(\mathfrak{M})$ *then* $\vdash_{\mathfrak{M}} \varphi$*.*

Proof. The proof is by induction on the construction of φ.

First we prove that if t is any closed $\mathscr{L}_{\mathfrak{M}}$-term and

$$\mathfrak{M} \vDash t = \mathsf{c}_a \quad \text{then} \quad \vdash_{\mathfrak{M}} t = \mathsf{c}_a.$$

This is done by induction on the construction of t. The base case, when t is a constant, follows from Axiom3. We prove the case when t is $r + s$ for some closed terms r and s. Let b, $d \in \mathfrak{M}$ be such that

$$\mathfrak{M} \vDash r = \mathsf{c}_b \wedge s = \mathsf{c}_d \wedge b + d = a.$$

By the induction hypothesis $\vdash_{\mathfrak{M}} r = \mathsf{c}_b$ and $\vdash_{\mathfrak{M}} s = \mathsf{c}_d$, so by Axiom7 $\vdash_{\mathfrak{M}} t = \mathsf{c}_b + \mathsf{c}_d$ and by Axiom10 and Axiom5 $\vdash_{\mathfrak{M}} t = \mathsf{c}_a$. The other cases, when t is $\mathsf{S}(r)$ or $r \cdot s$ are similar.

Now we prove, by induction on the construction of φ, that

$$\begin{array}{llll} \text{if} & \mathfrak{M} \vDash \varphi & \text{then} & \vdash_{\mathfrak{M}} \varphi \qquad \text{and} \\ \text{if} & \mathfrak{M} \nvDash \varphi & \text{then} & \vdash_{\mathfrak{M}} \neg\varphi. \end{array}$$

For the base case, when φ is an atomic sentence, assume first that $\mathfrak{M} \vDash t = r$. Then, clearly, $\mathfrak{M} \vDash t = \mathsf{c}_a \wedge r = \mathsf{c}_a$ for some $a \in \mathfrak{M}$. By the fact proved above $\vdash_{\mathfrak{M}} t = \mathsf{c}_a$ and $\vdash_{\mathfrak{M}} r = \mathsf{c}_a$, and by using Axiom3 and Axiom4 by get $\vdash_{\mathfrak{M}} t = r$.

If $\mathfrak{M} \vDash t \neq r$ then $\mathfrak{M} \vDash t = \mathsf{c}_a \wedge r = \mathsf{c}_b$ for some $a \neq b \in \mathfrak{M}$. The formal deduction of $t \neq r$ from $t = \mathsf{c}_a$; $r = \mathsf{c}_b$ and $\mathsf{c}_a \neq \mathsf{c}_b$ is written out in Figure 3.3 as an illustration of a $\mathfrak{M}$-logic proof.

For the inductive step we only handle the case when φ is $\psi \vee \sigma$, the others are similar and easy. Assume $\mathfrak{M} \vDash \psi \vee \sigma$ then $\mathfrak{M} \vDash \psi$ or $\mathfrak{M} \vDash \sigma$ so by the induction hypothesis we have $\vdash_{\mathfrak{M}} \psi$ or $\vdash_{\mathfrak{M}} \sigma$. In either case $\vdash_{\mathfrak{M}} \psi \vee \sigma$ by $\vee$I1 or $\vee$I2.

On the other hand if $\mathfrak{M} \nvDash \psi \vee \sigma$ then $\mathfrak{M} \nvDash \psi$ and $\mathfrak{M} \nvDash \sigma$. Thus, by the induction hypothesis $\vdash_{\mathfrak{M}} \neg\psi$ and $\vdash_{\mathfrak{M}} \neg\sigma$. An application of $\vee$I3 yields $\vdash_{\mathfrak{M}} \neg(\psi \vee \sigma)$. $\square$

3.2.2 Soundness and completeness

If $\Gamma = \{ \gamma_1, \gamma_2, \ldots, \gamma_k \}$ is a finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, let $\bigvee \Gamma$ be

$$\gamma_1 \vee (\gamma_2 \vee (\ldots \vee \gamma_k))$$

$$
\begin{prooftree}
\AxiomC{Axiom5}
\UnaryInfC{$\mathsf{c}_a = \mathsf{c}_b, r \neq \mathsf{c}_b, \mathsf{c}_a \neq r$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b, r \neq \mathsf{c}_b, \mathsf{c}_a \neq r$}
\AxiomC{Axiom5}
\UnaryInfC{$t \neq r, \mathsf{c}_a = r, \mathsf{c}_a \neq t$}
\AxiomC{Axiom4}
\UnaryInfC{$\mathsf{c}_a = t, t \neq \mathsf{c}_a$}
\AxiomC{$t = \mathsf{c}_a$}
\BinaryInfC{$\mathsf{c}_a = t$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = t$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = r, \mathsf{c}_a = t$}
\BinaryInfC{$t \neq r, \mathsf{c}_a = r$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b, \mathsf{c}_a = r$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b, r \neq \mathsf{c}_b, \mathsf{c}_a = r$}
\BinaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b, r \neq \mathsf{c}_b$}
\AxiomC{$r = \mathsf{c}_b$}
\UnaryInfC{$t \neq r, r = \mathsf{c}_b$}
\UnaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b, r = \mathsf{c}_b$}
\BinaryInfC{$t \neq r, \mathsf{c}_a = \mathsf{c}_b$}
\AxiomC{Axiom2}
\UnaryInfC{$\mathsf{c}_a \neq \mathsf{c}_b$}
\UnaryInfC{$t \neq r, \mathsf{c}_a \neq \mathsf{c}_b$}
\BinaryInfC{$t \neq r$}
\end{prooftree}
$$

Figure 3.3: *A proof of* $t \neq r$ *from the hypothesis* $t = \mathsf{c}_a$*;* $r = \mathsf{c}_b$ *and* $a \neq b$*.*

if Γ is nonempty and $0 \neq 0$ if Γ is empty.[2] If Σ is a satisfaction class then $\bigvee \Gamma \in \Sigma$ iff $\gamma_i \in \Sigma$ for some $1 \leqslant i \leqslant k$. Please note that this is independent on the numbering of the set Γ.

Proposition 3.20 (Soundness of $\mathfrak{M}$-logic). *If $\vdash_{\mathfrak{M}} \Gamma$ and Σ is a satisfaction class then $\bigvee \Gamma \in \Sigma$.*

Proof. The proof is by induction on the height of the proof p of Γ. If $|p| = 0$ then Γ is an axiom:

Axiom1: If Γ is $\psi, \neg\psi$, it is clear that $\psi \vee \neg\psi \in \Sigma$ due to the definition of satisfaction classes.

Axiom2: If $a \neq b$ then $\mathsf{c}_a \neq \mathsf{c}_b \in \Sigma$ by (3.2).

Axiom3: If Γ is $t = t$ then $t = t \in \Sigma$ by (3.3).

Axiom4: This is (3.4).

Axiom5: This is (3.5).

Axiom6: If Γ is $t \neq r, \mathsf{S}(t) = \mathsf{S}(r)$ we have to prove that $t \neq r \vee \mathsf{S}(t) = \mathsf{S}(r) \in \Sigma$ or equivalent that if $t = r \in \Sigma$ then $\mathsf{S}(t) = \mathsf{S}(r) \in \Sigma$. If $t = r \in \Sigma$ let a be such that $t = \mathsf{c}_a \in \Sigma$ and $r = \mathsf{c}_a \in \Sigma$. We have $\mathsf{S}(t) = \mathsf{c}_b \in \Sigma$ and $\mathsf{S}(r) = \mathsf{c}_b \in \Sigma$, where $b = \mathsf{S}^{\mathfrak{M}}(a)$, by (3.6). Thus, by (3.3)-(3.5), $\mathsf{S}(t) = \mathsf{S}(r) \in \Sigma$.

Axiom7: As above but use (3.7) instead of (3.6).

Axiom8: As above but use (3.8) instead of (3.6).

Axiom9, Axiom10, Axiom11: Directly from Proposition 3.4.

Axiom12: $\exists \mathsf{v}_0(t = \mathsf{v}_0) \in \Sigma$ since by (3.3) we have $t = t \in \Sigma$ and by (3.9) there exists $a \in \mathfrak{M}$ such that $t = \mathsf{c}_a \in \Sigma$ and so $\exists \mathsf{v}_0(t = \mathsf{v}_0) \in \Sigma$ by (3.12).

For the inductive step we prove that all the inference rules are sound in the sense that if all disjunctions of the premises of an inference rule are in in a satisfaction class then so is the disjunction of the conclusion. This is easy to see for the first six rules. We prove it only for $\exists$I and $\mathfrak{M}$-rule.

$\exists$I: Suppose $\bigvee \Lambda \vee \psi[\mathsf{c}_a/\mathsf{v}_i] \in \Sigma$; then we can assume that

$$\psi[\mathsf{c}_a/\mathsf{v}_i] \in \Sigma$$

since otherwise $\bigvee \Lambda \in \Sigma$ and $\bigvee \Lambda \vee \exists \mathsf{v}_i\, \psi \in \Sigma$. Thus, by (3.1), $\exists \mathsf{v}_i\, \psi \in \Sigma$ and so $\bigvee \Lambda \vee \exists \mathsf{v}_i\, \psi \in \Sigma$.

$\mathfrak{M}$-rule: Suppose $\bigvee \Lambda \vee \neg\psi[\mathsf{c}_a/\mathsf{v}_i] \in \Sigma$ for all $a \in \mathfrak{M}$. We can assume

$$\neg\psi[\mathsf{c}_a/\mathsf{v}_i] \in \Sigma$$

for all $a \in \mathfrak{M}$. If $\neg\exists \mathsf{v}_i\, \psi \notin \Sigma$ then $\exists \mathsf{v}_i\, \psi \in \Sigma$ and then $\psi[\mathsf{c}_a/\mathsf{v}_i] \in \Sigma$ for some $a \in \mathfrak{M}$ by (3.1), and we have a contradiction. $\square$

Remark 3.21. There is a stronger version of the proposition: If $\Lambda \vdash_{\mathfrak{M}} \Gamma$ and Σ is a satisfaction class extending the set Λ, i.e., $\Lambda \subseteq \Sigma$, then $\bigvee \Gamma \in \Sigma$.

[2] We will later redefine $\bigvee \Gamma$ so that it does not depend on the numbering of the set by choosing a canonical numbering.

Suppose $\mathfrak{M}$ is countable and the set Λ is a set of ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentences which is consistent (e.g., if $\mathfrak{M}$-logic is consistent itself we can choose Λ to be the empty set). We will construct a satisfaction class including Λ. The construction is very much as the completeness theorem for first-order logic.

Let $\{\ \varphi_i\ \}_{i=1}^{\infty}$ be an enumeration of all ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentences and let $\Gamma_0 = \Lambda$. Define the sequence $\{\ \Gamma_i\ \}_{i=0}^{\infty}$ of consistent sets of ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentences recursively as follows.

We know that if Γ_i is consistent then either Γ_i, φ_{i+1} or $\Gamma_i, \neg\varphi_{i+1}$ is consistent. Choose Γ_{i+1} to be the one which is consistent with the extra condition that if φ_{i+1} is $\exists \mathsf{v}_j\, \psi$ for some ${}^{*}\mathscr{L}_{\mathfrak{M}}$-formula ψ and we choose Γ_{i+1} to be $\Gamma_i, \exists \mathsf{v}_j\, \psi$ then we also put $\psi[\mathsf{c}_a/\mathsf{v}_j]$ in Γ_{i+1} in such a way that this new set is consistent. We can always do this, because otherwise we would have

$$\Gamma_i, \exists \mathsf{v}_j\, \psi \vdash_{\mathfrak{M}} \neg\psi[\mathsf{c}_a/\mathsf{v}_j]$$

for all $a \in \mathfrak{M}$, but then

$$\Gamma_i, \exists \mathsf{v}_j\, \psi \vdash_{\mathfrak{M}} \neg\exists \mathsf{v}_j\, \psi$$

by $\mathfrak{M}$-rule so by Cut $\Gamma_i, \exists \mathsf{v}_i\, \psi \vdash_{\mathfrak{M}} \emptyset$ contradicting the fact that $\Gamma_i, \exists \mathsf{v}_i\, \psi$ is consistent.

Let

$$\Gamma_\infty = \bigcup \Gamma_i.$$

Lemma 3.22. Γ_∞ *is a maximally consistent set of* ${}^{*}\mathscr{L}_{\mathfrak{M}}$*-sentences.*

Proof. It is clearly maximal from the construction: for every sentence φ we have $\varphi \in \Gamma_\infty$ or $\neg\varphi \in \Gamma_\infty$. To prove the consistency of Γ_∞ we first notice that by the construction there is no sentence φ such that $\varphi \in \Gamma_\infty$ and $\neg\varphi \in \Gamma_\infty$. Then we observe that for all rules with finitely many premises if the premises are in Γ_∞ then the conclusion is too. This is easy to see since all the premises are in some Γ_n. Assume now that $\neg\varphi[\mathsf{c}_a/\mathsf{v}_i] \in \Gamma_\infty$ for all $a \in \mathfrak{M}$ and that $\exists \mathsf{v}_0\, \varphi \in \Gamma_\infty$, by the construction we have $\varphi[\mathsf{c}_a/\mathsf{v}_i] \in \Gamma_\infty$ for some a which is a contradiction. We have proved that Γ_∞ is closed under all the inference rules so since there is no ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentence $\varphi \in \Gamma_\infty$ such that $\neg\varphi \in \Gamma_\infty$, Γ_∞ is consistent. $\square$

Proposition 3.23. Σ *is a maximally consistent set of* ${}^{*}\mathscr{L}_{\mathfrak{M}}$*-sentences iff* Σ *is a satisfaction class.*

Proof. If Σ is a satisfaction class we can prove the consistency and the maximality very much as in Lemma 3.22. We prove the converse.

Assume Σ is a maximally consistent set of ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentences. We will use Proposition 3.6 to prove that Σ is a satisfaction class. First we have to prove that $\sim$ is an equivalence relation; Axiom3, Axiom4 and Axiom5 takes care of that. Axiom6, Axiom7 and Axiom8 takes care of the well-definability of $\mathfrak{M}_\Sigma$. The canonical map $f : \mathfrak{M} \to \mathfrak{M}_\Sigma$ is an embedding by Axiom2, Axiom9, Axiom10 and Axiom11, f is surjective by Axiom12. $\square$

Adding up the results in this section we get the following:

Proposition 3.24. *If* Λ *is a set of* ${}^{*}\mathscr{L}_{\mathfrak{M}}$*-sentences then there is a satisfaction class* Σ *such that* $\Lambda \subseteq \Sigma$ *iff* $\Lambda \nvdash_{\mathfrak{M}} \emptyset$.

Remark 3.25. We can define $\Lambda \models_{\mathfrak{M}} \Gamma$ for sets of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences Γ and Λ, where Γ is finite, to hold iff $\bigvee \Gamma$ is included in every satisfaction class extending Λ. Then Proposition 3.24 can be reformulated as

$$\Lambda \vdash_{\mathfrak{M}} \Gamma \quad \text{iff} \quad \Lambda \models_{\mathfrak{M}} \Gamma.$$

Finally, we add a small remark for the confused reader.

Remark 3.26. Even though by definition a consistent set Γ is maximally consistent in a logic if for every sentence or formula $\varphi \notin \Gamma$, the set $\Gamma \cup \{ \varphi \}$ is inconsistent; in a logic where the deduction theorem holds a consistent set Γ is maximally consistent iff for all sentences or formulas φ either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.

3.2.3 The height of proofs

The next result is about recursively saturated models. This is the only time we use the recursive saturation in this chapter. It tells us that in recursively saturated models we can prove everything provable in $\mathfrak{M}$-logic with proofs of finite height.

Proposition 3.27. *If Λ is an $\mathscr{L}_{\mathfrak{M}}$-definable set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, $\Lambda \vdash_{\mathfrak{M}} \Gamma$ and $\mathfrak{M}$ is recursively saturated then $\Lambda \vdash^{\omega}_{\mathfrak{M}} \Gamma$.*

Proof. Let $\lambda(x)$ define Λ. We will recursively define formulas $\mathtt{Pf}_k(x)$, $k \geqslant 1$ such that

$$\mathfrak{M} \models \mathtt{Pf}_k(a) \quad \text{iff} \quad \Lambda \vdash^{k}_{\mathfrak{M}} \Gamma,$$

where a is a code for the finite set Γ.

Let $\mathtt{Pf}_1(x)$ be

$$\begin{aligned}
&\forall z\big(z \in x \rightarrow \mathtt{Sent}(z)\big)\\
&\wedge \big[\exists y\big(\mathtt{Sent}(y) \wedge x = \{ y \} \wedge \lambda(y)\big)\\
&\vee \exists y\big(\mathtt{Sent}(y) \wedge x = \{ y, \ulcorner\neg y\urcorner \}\big)\\
&\vee \exists y\big(x = \{ \ulcorner \mathsf{c}_y = \mathsf{c}_y \urcorner \}\big)\\
&\vee \exists y,z\big(x = \{ \ulcorner \mathsf{c}_y \neq \mathsf{c}_z \urcorner \} \wedge y \neq z\big)\\
&\vee \exists y\big(\mathtt{ClTerm}(y) \wedge x = \{ \ulcorner y = y \urcorner \}\big)\\
&\vee \exists y,z\big(\mathtt{ClTerm}(y) \wedge \mathtt{ClTerm}(z) \wedge x = \{ \ulcorner y \neq z \urcorner, \ulcorner z = y \urcorner \}\big)\\
&\vee \exists y,z,w\big(\mathtt{ClTerm}(y) \wedge \mathtt{ClTerm}(w) \wedge \mathtt{ClTerm}(z) \wedge\\
&\qquad x = \{ \ulcorner y \neq z \urcorner, \ulcorner z \neq w \urcorner, \ulcorner y = w \urcorner \}\big)\\
&\vee \exists y,z\big(\mathtt{ClTerm}(y) \wedge x = \Big\{ \ulcorner y \neq \mathsf{c}_z \urcorner, \ulcorner \mathsf{S}(y) = \mathsf{S}(\mathsf{c}_z) \urcorner \Big\}\big)\\
&\vee \exists y,z,w,v\big(\mathtt{ClTerm}(y) \wedge \mathtt{ClTerm}(z) \wedge\\
&\qquad x = \{ \ulcorner y \neq \mathsf{c}_w \urcorner, \ulcorner z \neq \mathsf{c}_v \urcorner, \ulcorner y + z = \mathsf{c}_w + \mathsf{c}_v \urcorner \}\big)\\
&\vee \exists y,z,w,v\big(\mathtt{ClTerm}(y) \wedge \mathtt{ClTerm}(z) \wedge\\
&\qquad x = \{ \ulcorner y \neq \mathsf{c}_w \urcorner, \ulcorner z \neq \mathsf{c}_v \urcorner, \ulcorner y \cdot z = \mathsf{c}_w \cdot \mathsf{c}_v \urcorner \}\big)\\
&\vee \exists y\big(\mathtt{ClTerm}(y) \wedge x = \Big\{ \ulcorner \exists \mathsf{v}_0 (y = \mathsf{v}_0) \urcorner \Big\}\big)\big]
\end{aligned}$$

And let $\mathtt{Pf}_{k+1}(x)$ $(k \geqslant 1)$ be

$$\begin{aligned}
&\exists y, z\big(\mathtt{Sent}(z) \wedge \mathtt{Pf}_k(y) \wedge x = y \cup \{\, z \,\}\big)\\
&\vee \exists y, z, w\big(\mathtt{Sent}(z) \wedge \mathtt{Sent}(w) \wedge \mathtt{Pf}_k(y \cup \{\, z \,\}) \wedge x = y \cup \{\, \ulcorner z \vee w \urcorner \,\}\big)\\
&\vee \exists y, z, w\big(\mathtt{Sent}(z) \wedge \mathtt{Sent}(w) \wedge \mathtt{Pf}_k(y \cup \{\, w \,\}) \wedge x = y \cup \{\, \ulcorner z \vee w \urcorner \,\}\big)\\
&\vee \exists y, z, w\big(\mathtt{Sent}(z) \wedge \mathtt{Sent}(w) \wedge \mathtt{Pf}_k(y \cup \{\, \ulcorner \neg z \urcorner \,\}) \wedge \mathtt{Pf}_k(y \cup \{\, \ulcorner \neg w \urcorner \,\}) \wedge\\
&\qquad x = y \cup \Big\{\, \ulcorner \neg(z \vee w) \urcorner \,\Big\}\big)\\
&\vee \exists y, z\big(\mathtt{Sent}(z) \wedge \mathtt{Pf}_k(y \cup \{\, z \,\}) \wedge x = y \cup \{\, \ulcorner \neg\neg z \urcorner \,\}\big)\\
&\vee \exists z\big(\mathtt{Sent}(z) \wedge \mathtt{Pf}_k(x \cup \{\, z \,\}) \wedge \mathtt{Pf}_k(x \cup \{\, \ulcorner \neg z \urcorner \,\})\big)\\
&\vee \exists y, z, w, i\big(\mathtt{Sent}(\ulcorner \exists \mathsf{v}_i\, z \urcorner) \wedge \mathtt{Pf}_k(y \cup \Big\{\, \ulcorner z[\mathsf{c}_w/\mathsf{v}_i] \urcorner \,\Big\}) \wedge x = y \cup \{\, \ulcorner \exists \mathsf{v}_i\, z \urcorner \,\}\big)\\
&\vee \exists y, z, i(\mathtt{Sent}(\ulcorner \neg\exists \mathsf{v}_i\, z \urcorner) \wedge \forall w\, \mathtt{Pf}_k(y \cup \Big\{\, \ulcorner \neg z[\mathsf{c}_w/\mathsf{v}_i] \urcorner \,\Big\}) \wedge x = y \cup \{\, \ulcorner \neg\exists \mathsf{v}_i\, z \urcorner \,\}).
\end{aligned}$$

Observe that if $a \in \mathfrak{M}$ is code for a finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences Γ and $\mathfrak{M} \models \mathtt{Pf}_k(a)$ then $|\Gamma| \leqslant k + 2$.

Suppose the lemma is false and let

$$A = \{\, \Gamma \mid \Lambda \vdash_{\mathfrak{M}} \Gamma \text{ and } \Lambda \nvdash^{\omega}_{\mathfrak{M}} \Gamma \,\}$$

and

$$B = \{\, p \mid p \text{ is a proof from } \Lambda \text{ of } \Gamma \text{ for some } \Gamma \in A \,\}\,.$$

Let $p \in B$ be of smallest height, i.e., such that if $q \in B$ then $|p| \leqslant |q|$. All the subproofs of p must be of finite height, so it is clear that $|p| = \omega$ and that the last inference rule in p is $\mathfrak{M}$-rule:

$$\frac{\ldots\ \Delta, \neg\varphi[\mathsf{c}_a/\mathsf{v}_i]\ \ldots\ {}_{a \in \mathfrak{M}}}{\Delta, \neg\exists \mathsf{v}_i\, \varphi}$$

Let $a \in \mathfrak{M}$ be a code for Δ and define

$$q(x) = \Big\{\, \neg\, \mathtt{Pf}_k\big(a \cup \Big\{\, \ulcorner \neg\varphi[\mathsf{c}_x/\mathsf{v}_i] \urcorner \,\Big\}\big) \,\Big|\, k \geqslant 1 \,\Big\}\,.$$

It should be clear that $q(x)$ is a recursive type so, by recursive saturation, it is realized by some $b \in \mathfrak{M}$. We therefore have

$$\nvdash^{\omega}_{\mathfrak{M}} \Delta, \neg\varphi[\mathsf{c}_b/\mathsf{v}_i],$$

which contradicts the fact that the height of p is ω. □

Remark 3.28. In Proposition 3.27 we do not need Λ to be definable, we only need the expanded structure $\langle \mathfrak{M}, \Lambda \rangle$ to be recursively saturated for the proof to work.

3.3 The inconsistency of $\mathfrak{M}$-logic

In this section we prove that if $\mathfrak{M}$ admits a satisfaction class then $\mathfrak{M}$ is recursively saturated (without any restriction on the cardinality of $\mathfrak{M}$). We will also mention a strengthening of this result by Smith.

Please do compare the next theorem (and proof) with Theorem 3.11.

Theorem 3.29 ([Lac81]). *Let $\mathfrak{M}$ be an arbitrary nonstandard model of* PA *admitting a satisfaction class then $\mathfrak{M}$ is recursively saturated.*

Proof. Let Σ be a satisfaction class on $\mathfrak{M}$ and assume $\mathfrak{M}$ is not recursively saturated. Let

$$\{\, \varphi_i(x) \,\}_{i\in\omega}$$

be a non-realized recursive type. We can assume that

$$\begin{aligned}
\mathfrak{M} &\models \forall x\big(\varphi_{i+1}(x) \to \varphi_i(x)\big),\\
\mathfrak{M} &\models \exists x\big(\neg\varphi_{i+1}(x) \wedge \varphi_i(x)\big) \quad \text{and}\\
\mathfrak{M} &\models \forall x\, \varphi_0(x),
\end{aligned}$$

for all $i \in \omega$. If not, we can replace the formulas φ_i by φ'_i where

$$\begin{aligned}
\varphi'_0(x) &\quad \text{is} \quad x = x \qquad \text{and}\\
\varphi'_{i+1}(x) &\quad \text{is} \quad \varphi'_i(x) \wedge \varphi_i(x) \wedge \exists y{<}x\, \varphi'_i(y).
\end{aligned}$$

Let

$$\alpha_i(x) \text{ be } \neg\varphi_{i+1}(x) \wedge \varphi_i(x) \quad \text{for all } i \in \omega$$

and

$$A_i = \{\, a \in \mathfrak{M} \mid \mathfrak{M} \models \alpha_i(a) \,\}.$$

It should be clear that $\{\, A_i \,\}_{i\in\omega}$ is a partition of $\mathfrak{M}$.

We will recursively define a sequence $\{\, \beta_i(x) \,\}_{i\leqslant\nu}$ of formulas. Let $\beta_0(x)$ be $\alpha_0(x)$ and if $\beta_i(x)$ is defined let θ_{ij} be

$$\exists x\big(\beta_i(x) \wedge \alpha_j(x)\big)$$

and $\beta_{i+1}(x)$ be

$$\begin{aligned}
\big(\neg\exists y \beta_i(y) \wedge \alpha_0(x)\big) \vee \Big(\exists y\, \beta_i(y) \wedge \big(\theta_{i0} \wedge \alpha_1(x)\big) \vee \big[\neg\theta_{i0}&\\
\wedge \big(\theta_{i1} \wedge \alpha_2(x)\big) \vee \big[\neg\theta_{i1}&\\
\wedge \big(\theta_{i2} \wedge \alpha_3(x)\big) \vee \big[\neg\theta_{i2}&\\
\vdots\quad&\\
\wedge \big(\theta_{ii} \wedge \alpha_{i+1}(x)\big)\cdots\big]\big]\big]\Big).&
\end{aligned}$$

The sequence $\beta_i(x)$ is recursive and therefore coded in $\mathfrak{M}$ and so extendable to a nonstandard $\nu \in \mathfrak{M} \setminus \omega$, so that the recursive definition holds for all $i < \nu$. Let

$$B_i = \{\, a \in \mathfrak{M} \mid \beta_i[\mathsf{c}_a/\mathsf{v}_j] \in \Sigma \,\}$$

for all $i \leqslant \nu$, where v_j is the free variable in $\beta_i(x)$.

The idea of the definitions of the β_is is that

$$B_{i+1} = \begin{cases} A_0 & \text{if } B_i = \emptyset \\ A_{k+1} & \text{otherwise, where } k = (\mu n{\in}\omega)\, B_i \cap A_n \neq \emptyset. \end{cases}$$

Since a satisfaction class is able to "look" finitely deep into a formula we can prove the following properties of the sequence $\{\, B_i \,\}_{i \leqslant \nu}$:

$$B_i = A_k \quad \Rightarrow \quad B_{i+1} = A_{k+1} \qquad \text{for all } i \leqslant \nu \text{ and } k \in \omega \tag{3.13}$$

by the recursive definition of $\beta_i(x)$, and

$$\forall i{\leqslant}\nu\, \exists k{\in}\omega\, B_i = A_k \tag{3.14}$$

since $B_0 = A_0$ and if $i > 0$ then either B_{i-1} is the empty set, in which case B_i is A_0, or B_{i-1} is not empty and then there is a least $k \in \omega$ such that B_{i-1} intersects A_k (since $\{\, A_i \,\}_{i \in \omega}$ is a partition of $\mathfrak{M}$) and then B_i is A_{k+1}.

Let us now finally define the (external) function $f : I_{<a} \to \omega$ such that $f(i) = k$ iff $B_i = A_k$. By the property (3.14) this is a total function and by property (3.13) $f(i+1) = f(i) + 1$, so the sequence

$$f(\nu) > f(\nu - 1) > f(\nu - 2) > \dots$$

is a strictly decreasing infinite sequence of natural numbers, which contradicts the well-ordering of the natural numbers. Hence there could not be a non-realized recursive type and therefore $\mathfrak{M}$ is recursively saturated. □

There is a somewhat stronger version of the theorem:

Theorem 3.30 ([Smi84]). *If $\mathfrak{M}$-logic is consistent then $\mathfrak{M}$ is recursively saturated.*

Proof. The proof is by modifying the proof of Lachlan's result; defining the sets B_i by provability in $\mathfrak{M}$-logic instead of by satisfaction classes, see [Smi84] for the details. □

Please observe that when $\mathfrak{M}$ is countable $\mathfrak{M}$ admits a satisfaction class precisely when $\mathfrak{M}$-logic is consistent, so the result of Smith is a strengthening of Lachlan's result only when $\mathfrak{M}$ is uncountable.

3.4 The consistency of $\mathfrak{M}$-logic

3.4.1 Template logic

In Section 3.2 we proved that $\mathfrak{M}$-logic is consistent iff we can find a satisfaction class for the model (assuming $\mathfrak{M}$ is countable). In this section we prove that for any recursively saturated model of $\mathtt{PA}$ $\mathfrak{M}$-logic is consistent.

Together with the results in Section 3.3 above we get that a countable non-standard model of $\mathtt{PA}$ admits a satisfaction class iff it is recursively saturated.

The question is; if $\mathfrak{M}$ is recursively saturated how do we prove the consistency of $\mathfrak{M}$-logic? Proving consistency of a logic can be done, mainly, in two different ways; the proof theoretic way, by a cut-elimination theorem, or the model theoretic way, by a soundness theorem. We will use the model theoretic approach and prove a soundness theorem. For this we will define a new kind of logic. The idea is really easy; instead of studying nonstandard formulas we are going to replace some subformulas and subterms in the nonstandard formula with templates, and in this way study formulas of finite depth.

We will call this logic template logic, it is first-order logic with template symbols added. Each ${}^*\mathscr{L}_{\mathfrak{M}}$-formula and ${}^*\mathscr{L}_{\mathfrak{M}}$-term has a corresponding template symbol. These symbols may be looked on as predicates and functions of nonstandard finite arity. For example, the template symbol corresponding to the ${}^*\mathscr{L}_{\mathfrak{M}}$-formula

$$\mathsf{v}_0 = \mathsf{v}_0 \vee \mathsf{v}_1 = \mathsf{v}_1 \vee \ldots \vee \mathsf{v}_a = \mathsf{v}_a$$

could be treated as a predicate of arity a, even if $a > \omega$.

Let $\mathscr{L}_{\mathtt{T}}$ be the language

$$\mathscr{L}_{\mathfrak{M}} \cup \{\, \widehat{\varphi} \,|\, \varphi \text{ is a } {}^*\mathscr{L}_{\mathfrak{M}}\text{-formula} \,\} \cup \left\{\, \widehat{t} \,\middle|\, t \text{ is a } {}^*\mathscr{L}_{\mathfrak{M}}\text{-terms} \,\right\}$$

where $\widehat{\varphi}$ is treated as 0-ary relational symbols (i.e., propositional variables) and $\widehat{t}$ is treated as constant symbols when building up terms and formulas. For technical reasons we include the variables $\{\, \mathsf{v}_a \,|\, a \in \mathfrak{M} \,\}$ in the language $\mathscr{L}_{\mathtt{T}}$.

The formulas and terms of template logic are defined in the usual way with template symbols being 0-ary. The free variables of a term is defined in a non-standard way as follows

$$\begin{aligned}
\mathtt{FV}(\mathsf{c}_a) &=_{\mathrm{df}} \emptyset \\
\mathtt{FV}(\mathsf{v}_i) &=_{\mathrm{df}} \{\, \ulcorner \mathsf{v}_i \urcorner \,\} \\
\mathtt{FV}(\widehat{t}\,) &=_{\mathrm{df}} \{\, a \in \mathfrak{M} \,|\, \mathfrak{M} \vDash a \in \mathtt{FV}(t) \,\} \\
\mathtt{FV}(\mathsf{S}(r)) &=_{\mathrm{df}} \mathtt{FV}(r) \\
\mathtt{FV}(r + s) &=_{\mathrm{df}} \mathtt{FV}(r) \cup \mathtt{FV}(s) \\
\mathtt{FV}(r \cdot s) &=_{\mathrm{df}} \mathtt{FV}(r) \cup \mathtt{FV}(s)
\end{aligned}$$

Observe that in the third clause $\mathtt{FV}$ has two different meanings, in the second appearance it is the function definable in $\mathtt{PA}$. And of a formula as

$$\begin{aligned} \mathtt{FV}(r = s) &=_{\text{df}} \mathtt{FV}(r) \cup \mathtt{FV}(s) \\ \mathtt{FV}(\widehat{\varphi}) &=_{\text{df}} \{\, a \in \mathfrak{M} \mid \mathfrak{M} \vDash a \in \mathtt{FV}(\varphi) \,\} \\ \mathtt{FV}(\neg\gamma) &=_{\text{df}} \mathtt{FV}(\gamma) \\ \mathtt{FV}(\gamma \vee \delta) &=_{\text{df}} \mathtt{FV}(\gamma) \cup \mathtt{FV}(\delta) \\ \mathtt{FV}(\exists \mathsf{v}_i\, \gamma) &=_{\text{df}} \mathtt{FV}(\gamma) \setminus \{\, \ulcorner \mathsf{v}_i \urcorner \,\} \end{aligned}$$

We call a formula γ a *sentence* if $\mathtt{FV}(\gamma) = \emptyset$ and a term t *closed* if $\mathtt{FV}(t) = \emptyset$.

Substitution is also defined in a nonstandard way; for terms

$$\begin{array}{rcl} \mathsf{c}_b[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \mathsf{c}_b \\ \mathsf{v}_j[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \begin{cases} \mathsf{c}_a & \text{if } i = j \\ \mathsf{v}_j & \text{otherwise} \end{cases} \\ \widehat{t}\,[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \widehat{t[\mathsf{c}_a/\mathsf{v}_i]} \\ \mathsf{S}(r)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \mathsf{S}(r[\mathsf{c}_a/\mathsf{v}_i]) \\ (r+s)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & r[\mathsf{c}_a/\mathsf{v}_i] + s[\mathsf{c}_a/\mathsf{v}_i] \\ (r \cdot s)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & r[\mathsf{c}_a/\mathsf{v}_i] \cdot s[\mathsf{c}_a/\mathsf{v}_i] \end{array}$$

and for formulas

$$\begin{array}{rcl} (r = s)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & r[\mathsf{c}_a/\mathsf{v}_i] = s[\mathsf{c}_a/\mathsf{v}_i] \\ \widehat{\varphi}[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \widehat{\varphi[\mathsf{c}_a/\mathsf{v}_i]} \\ (\neg\gamma)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \neg(\gamma[\mathsf{c}_a/\mathsf{v}_i]) \\ (\gamma \vee \delta)[\mathsf{c}_a/\mathsf{v}_i] & \text{is} & \gamma[\mathsf{c}_a/\mathsf{v}_i] \vee \delta[\mathsf{c}_a/\mathsf{v}_i] \\ \exists \mathsf{v}_i\, \gamma[\mathsf{c}_a/\mathsf{v}_j] & \text{is} & \begin{cases} \exists \mathsf{v}_i(\gamma[\mathsf{c}_a/\mathsf{v}_j]) & \text{if } i \neq j \\ \exists \mathsf{v}_i\, \gamma & \text{otherwise.} \end{cases} \end{array}$$

Definition 3.31. If φ is a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula and $a \in \mathfrak{M}$ then $\varphi[a/\mathsf{v}]$ is

$$\varphi[\mathsf{c}_{[a]_{i_1}-1}, \ldots, \mathsf{c}_{[a]_{i_k}-1}/\mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_k}]$$

where $\{\, i_1, \ldots, i_k \,\} = \{\, i \in \mathfrak{M} \mid [a]_i \neq 0 \,\}$. We define $t[a/\mathsf{v}]$, for terms t, in the obvious similar way.

Definition 3.32. The relation (between ${}^*\mathscr{L}_{\mathfrak{M}}$-formulas) $\varphi \cong \psi$ holds if there exists a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula γ and a, $b \in \mathfrak{M}$ such that $\gamma[a/\mathsf{v}]$ is φ and $\gamma[b/\mathsf{v}]$ is ψ. The relation between ${}^*\mathscr{L}_{\mathfrak{M}}$-terms $t \cong r$ holds if there exists a ${}^*\mathscr{L}_{\mathfrak{M}}$-term s and a, $b \in \mathfrak{M}$ such that $s[a/\mathsf{v}]$ is t and $s[b/\mathsf{v}]$ is r.

Proposition 3.33. *The relation $\cong$ is an equivalence relation on the set of ${}^*\mathscr{L}_{\mathfrak{M}}$-formulas and ${}^*\mathscr{L}_{\mathfrak{M}}$-terms.*

Proof. The reflexive and symmetric properties are trivial, only the transitive property involves some work. We prove the proposition for terms, the case with formulas is similar.

Let us assume that $t \cong r$ and $r \cong s$, and prove $t \cong s$. The proof is by induction on r inside PA; we use the inductive property of $\mathtt{Term}(x)$. If r is a constant or a variable then t and s are also either constants or variables. In any case $t \cong s$ since any constant is related to any constant or variable and a variable is related to a variable iff they are equal.

For the inductive step the case when r is of the form $\mathsf{S}(r')$ is easy and the two cases when r is $r_1 + r_2$ or $r_1 \cdot r_2$ are similar, therefore we only handle the case when r is $r_1 + r_2$.

It is easy to see that t and s also are of this form, i.e., t is $t_1 + t_2$ and s is $s_1 + s_2$. Let p, q, a_t, a_r, b_r, and b_s be such that

$$t = p[a_t/\mathsf{v}], r = p[a_r/\mathsf{v}], r = q[b_r/\mathsf{v}] \quad \text{and} \quad s = q[b_s/\mathsf{v}]$$

and let p be $p_1 + p_2$ and q be $q_1 + q_2$. Clearly,

$$t_i = p_i[a_t/\mathsf{v}], r_i = p_i[a_r/\mathsf{v}], r_i = q_i[b_r/\mathsf{v}] \quad \text{and} \quad s_i = q_i[b_s/\mathsf{v}]$$

for $i = 1, 2$, therefore $t_i \cong r_i$ and $r_i \cong s_i$. Thus, by the induction hypothesis $t_i \cong s_i$. Let o_i, a_{it} and a_{is} be such that

$$t_i = o_i[a_{it}/\mathsf{v}] \quad \text{and} \quad s_i = o_i[a_{is}/\mathsf{v}].$$

We have to "unify" a_{it} and a_{is} in such a way that the results, a_t and a_s, work for both is, i.e., such that $t_i = o'_i[a_t/\mathsf{v}]$ and $s_i = o'_i[a_s/\mathsf{v}]$, where o_i are some new terms. Then clearly $t = (o'_1 + o'_2)[a_t/\mathsf{v}]$ and $s = (o_1 + o_2)[a_s/\mathsf{v}]$ so $t \cong s$.

To make this happen let o'_1 be o_1 and let $i_0 \in \mathfrak{M}$ be bigger than all indices of free variables of o_1 and all is such that

$$[a_{1t}]_i \neq 0 \quad \text{or} \quad [a_{1s}]_i \neq 0.$$

Let o'_2 be as o_2 except that if v_i occurs as a free variable in o_2 and $[a_{2t}]_i \neq 0$ it is substituted by v_{i+i_0}. Finally, define a_t and a_s such that

$$\begin{aligned} [a_t]_i &= [a_{1t}]_i \quad \text{and} \\ [a_s]_i &= [a_{1s}]_i \end{aligned}$$

for all $i < i_0$ and

$$\begin{aligned} [a_t]_{i+i_0} &= [a_{2t}]_i \quad \text{and} \\ [a_s]_{i+i_0} &= [a_{2s}]_i \end{aligned}$$

for all $i \in \mathfrak{M}$. Clearly, $t_i = o'_i[a_t/\mathsf{v}]$ and $s_i = o'_i[a_s/\mathsf{v}]$ for $i = 1, 2$. □

So far, so good; we have a new logic to play with (even though we have not defined the axioms and inference rules yet). But how does template logic connect with $\mathfrak{M}$-logic, the object of study? For the connection to work we need a way to approximate a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula or a ${}^*\mathscr{L}_{\mathfrak{M}}$-term with a template formula or term. Given a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula you can, by replacing some of the subformulas and subterms with corresponding template symbols, e.g., replacing the subformula φ by $\widehat{\varphi}$, make an $\mathscr{L}_{\mathsf{T}}$-formula. This is the idea behind approximations.

If ψ is an $\mathscr{L}_{\mathsf{T}}$-formula and δ a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula we define $\mathscr{F}_{\delta}(\psi)$ to be ψ with all occurrences of symbols $\widehat{\varphi}$, for $\varphi \cong \delta$, replaced by

$$\begin{aligned} \widehat{t_1} = \widehat{t_2} \quad & \text{if } \varphi \text{ is } t_1 = t_2 \\ \neg\widehat{\gamma} \quad & \text{if } \varphi \text{ is } \neg\gamma \\ \widehat{\gamma} \vee \widehat{\sigma} \quad & \text{if } \varphi \text{ is } \gamma \vee \sigma \\ \exists \mathsf{v}_i\, \widehat{\gamma} \quad & \text{if } \varphi \text{ is } \exists \mathsf{v}_i\, \gamma \end{aligned}$$

and we define $\mathscr{F}_r(\psi)$, where r is a ${}^*\mathscr{L}_{\mathfrak{M}}$-term, to be ψ with all occurrences of symbols $\widehat{t}$, for $t \cong r$, replaced by

$$\begin{aligned} \mathsf{v}_i \quad & \text{if } t \text{ is } \mathsf{v}_i \\ \mathsf{c}_a \quad & \text{if } t \text{ is } \mathsf{c}_a \\ \mathsf{S}(\widehat{r}) \quad & \text{if } t \text{ is } \mathsf{S}(r) \\ \widehat{r} + \widehat{s} \quad & \text{if } t \text{ is } r + s \\ \widehat{r} \cdot \widehat{s} \quad & \text{if } t \text{ is } r \cdot s. \end{aligned}$$

Now, the definition of an approximation.

Definition 3.34. An $\mathscr{L}_{\mathsf{T}}$-formula (or $\mathscr{L}_{\mathsf{T}}$-term) ψ is an approximation of another $\mathscr{L}_{\mathsf{T}}$-formula (or $\mathscr{L}_{\mathsf{T}}$-term) δ if there exist ${}^*\mathscr{L}_{\mathfrak{M}}$-formulas or -terms $\tau_1, \ldots, \tau_k$ such that

$$\delta = \mathscr{F}_{\tau_k} \circ \mathscr{F}_{\tau_{k-1}} \circ \ldots \circ \mathscr{F}_{\tau_1}(\psi).$$

If

$$\mathscr{F} = \mathscr{F}_{\tau_k} \circ \mathscr{F}_{\tau_{k-1}} \circ \ldots \circ \mathscr{F}_{\tau_1}$$

we call $\mathscr{F}$ an approximating function and say that the length of $\mathscr{F}$, denoted $|\mathscr{F}|$, is k. We say that ψ is an approximation of a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula φ if ψ is $\mathscr{F}(\widehat{\varphi})$, for some approximating function $\mathscr{F}$.

We define approximations of finite sets of formulas by letting

$$\mathscr{F}(\Delta) = \{\, \mathscr{F}(\delta) \mid \delta \in \Delta \,\}.$$

Finally, for convenience, we define $\mathscr{F}(\varphi)$, where φ is a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence, to be $\mathscr{F}(\widehat{\varphi})$.

The formal proof system for template logic as just like the one for $\mathfrak{M}$-logic, but for completeness we write it down anyway. The axioms are

$$\gamma, \neg\gamma \tag{Axiom1$_t$}$$
$$\mathsf{c}_a \neq \mathsf{c}_b \quad \text{if } a \neq b \tag{Axiom2$_t$}$$
$$t = t \tag{Axiom3$_t$}$$
$$t \neq r, r = t \tag{Axiom4$_t$}$$
$$t \neq r, r \neq s, t = s \tag{Axiom5$_t$}$$
$$t \neq r, \mathsf{S}(t) = \mathsf{S}(r) \tag{Axiom6$_t$}$$
$$t \neq t', r \neq r', t + r = t' + r' \tag{Axiom7$_t$}$$
$$t \neq t', r \neq r', t \cdot r = t' \cdot r' \tag{Axiom8$_t$}$$
$$\mathsf{S}(\mathsf{c}_a) = \mathsf{c}_{\mathsf{S}(a)} \tag{Axiom9$_t$}$$
$$\mathsf{c}_a + \mathsf{c}_b = \mathsf{c}_{a+b} \tag{Axiom10$_t$}$$
$$\mathsf{c}_a \cdot \mathsf{c}_b = \mathsf{c}_{a \cdot b} \tag{Axiom11$_t$}$$
$$\exists \mathsf{v}_0 (t = \mathsf{v}_0) \tag{Axiom12$_t$}$$

where γ is an arbitrary $\mathscr{L}_{\mathsf{T}}$-sentence, and t, r and s are arbitrary closed $\mathscr{L}_{\mathsf{T}}$-terms. The inference rules are the following:

$$\frac{\Gamma}{\Gamma, \gamma} \tag{Weak$_t$}$$

$$\frac{\Gamma, \gamma}{\Gamma, \gamma \vee \delta} \tag{$\vee$I1$_t$}$$

$$\frac{\Gamma, \delta}{\Gamma, \gamma \vee \delta} \tag{$\vee$I2$_t$}$$

$$\frac{\Gamma, \neg\gamma \quad \Gamma, \neg\delta}{\Gamma, \neg(\gamma \vee \delta)} \tag{$\vee$I3$_t$}$$

$$\frac{\Gamma, \gamma}{\Gamma, \neg\neg\gamma} \tag{$\neg$I$_t$}$$

$$\frac{\Gamma, \gamma \quad \Gamma, \neg\gamma}{\Gamma} \tag{Cut$_t$}$$

$$\frac{\Gamma, \gamma[\mathsf{c}_a/\mathsf{v}_i]}{\Gamma, \exists \mathsf{v}_i\, \gamma} \tag{$\exists$I$_t$}$$

$$\frac{\ldots \ \Gamma, \neg\gamma[\mathsf{c}_a/\mathsf{v}_i] \ \ldots \ {}_{a \in \mathfrak{M}}}{\Gamma, \neg\exists \mathsf{v}_i\, \gamma} \tag{$\mathfrak{M}$-rule$_t$}$$

where Γ is an arbitrary finite set of $\mathscr{L}_{\mathsf{T}}$-sentences, γ, δ and $\exists \mathsf{v}_i\, \gamma$ are arbitrary $\mathscr{L}_{\mathsf{T}}$-sentences, t and r are arbitrary closed $\mathscr{L}_{\mathsf{T}}$-terms and $a \in \mathfrak{M}$.

Similar to $\mathfrak{M}$-logic $\models^p_{\mathsf{T}} \Delta$ means that p is a proof of Δ in template logic. All other definitions in $\mathfrak{M}$-logic transform almost verbatim to template logic, that is also the case for Lemma 3.16 and Proposition 3.18.

Please observe that if we are only studying standard formulas template logic extends $\mathfrak{M}$-logic, so any proof in $\mathfrak{M}$-logic using only standard formulas and terms is also a proof in template logic. On the other hand if we restrict template logic to formulas and terms without template symbols, then $\mathfrak{M}$-logic is an extension of template logic.

3.4.2 Some Technical Results

We need some more information on how the approximating functions work for later use.

Lemma 3.35. *If $\mathscr{F}$ is an approximating function and γ is an $\mathscr{L}_{\mathsf{T}}$-formula then $\mathscr{F}(\gamma[\mathsf{c}_a/\mathsf{v}_i]) = \mathscr{F}(\gamma)[\mathsf{c}_a/\mathsf{v}_i]$.*

Proof. Observe first that it is enough to prove the lemma for approximating functions $\mathscr{F} = \mathscr{F}_\tau$; the general result follows by "moving" one $\mathscr{F}_\tau$ at a time.

Since both substitution and approximating functions commute with the symbols $\neg$, $\vee$, $\exists$, $=$, S, $+$ and $\cdot$ we only have to check the base cases, i.e., when γ is a constant, variable or template symbol.

If γ is a constant or variable (or even more generally if γ does not contain any template symbols) we have

$$\mathscr{F}_\tau(\gamma[\mathsf{c}_a/\mathsf{v}_i]) = \gamma[\mathsf{c}_a/\mathsf{v}_i] = \mathscr{F}_\tau(\gamma)[\mathsf{c}_a/\mathsf{v}_i].$$

If γ is a template symbol, say $\widehat{\delta}$, it is clear that if $\tau \ncong \delta$ then

$$\mathscr{F}_\tau(\widehat{\sigma})[\mathsf{c}_a/\mathsf{v}_i] = \widehat{\sigma}[\mathsf{c}_a/\mathsf{v}_i] = \widehat{\sigma[\mathsf{c}_a/\mathsf{v}_i]} = \mathscr{F}_\tau(\widehat{\sigma[\mathsf{c}_a/\mathsf{v}_i]} = \mathscr{F}_\tau(\widehat{\sigma}[\mathsf{c}_a/\mathsf{v}_i]).$$

In the third equality we are using the fact that $\tau \ncong \delta[\mathsf{c}_a/\mathsf{v}_i]$, this is easy to see since if $\tau \cong \delta[\mathsf{c}_a/\mathsf{v}_i]$ then, since $\cong$ is a equivalence relation and $\delta[\mathsf{c}_a/\mathsf{v}_i] \cong \delta$, $\tau \cong \delta$.

Suppose ψ is $\neg\gamma$ and $\varphi \cong \psi$ then

$$\mathscr{F}_\varphi(\widehat{\psi})[\mathsf{c}_a/\mathsf{v}_i] = \neg\widehat{\gamma}[\mathsf{c}_a/\mathsf{v}_i] = \neg\widehat{\gamma[\mathsf{c}_a/\mathsf{v}_i]} = \mathscr{F}_\varphi(\widehat{\psi[\mathsf{c}_a/\mathsf{v}_i]}) = \mathscr{F}_\varphi(\widehat{\psi}[\mathsf{c}_a/\mathsf{v}_i]).$$

The case when ψ is $\gamma \vee \delta$ is treated in a similar way. Suppose that ψ is $\exists \mathsf{v}_j\, \gamma$ then

$$\mathscr{F}_\varphi(\widehat{\psi})[\mathsf{c}_a/\mathsf{v}_i] = (\exists \mathsf{v}_j\, \widehat{\gamma})[\mathsf{c}_a/\mathsf{v}_i] =$$
$$\begin{cases} \exists \mathsf{v}_j\, \widehat{\gamma[\mathsf{c}_a/\mathsf{v}_i]} = \mathscr{F}_\varphi(\widehat{\psi[\mathsf{c}_a/\mathsf{v}_i]}) = \mathscr{F}_\varphi(\widehat{\psi}[\mathsf{c}_a/\mathsf{v}_i]) & \text{if } i \neq j \\ \exists \mathsf{v}_j\, \widehat{\gamma} = \mathscr{F}_\varphi(\widehat{\psi}) = \mathscr{F}_\varphi(\widehat{\psi}[\mathsf{c}_a/\mathsf{v}_i]) & \text{otherwise.} \end{cases}$$

We also have to check the term cases. If the term is a composite term then it is handled just as the $\neg\gamma$ case. For the other cases we have

$$\mathscr{F}_{\mathsf{v}_i}(\widehat{\mathsf{v}_i})[\mathsf{c}_a/\mathsf{v}_j] = \mathsf{v}_i[\mathsf{c}_a/\mathsf{v}_j] =$$
$$\begin{cases} \mathsf{v}_i = \mathscr{F}_{\mathsf{v}_i}(\widehat{\mathsf{v}_i}[\mathsf{c}_a/\mathsf{v}_j]) & \text{if } i \neq j \\ \mathsf{c}_a = \mathscr{F}_{\mathsf{v}_i}(\widehat{\mathsf{c}_a}) = \mathscr{F}_{\mathsf{v}_i}(\widehat{\mathsf{v}_i}[\mathsf{c}_a/\mathsf{v}_j]) & \text{otherwise.} \end{cases}$$

and

$$\mathscr{F}_{\mathsf{v}_j}(\widehat{\mathsf{c}_b})[\mathsf{c}_b/\mathsf{v}_i] = \mathsf{c}_b = \mathscr{F}_{\mathsf{v}_j}(\widehat{\mathsf{c}_b}[\mathsf{c}_a/\mathsf{v}_i]). \qquad \square$$

Lemma 3.36. *If Δ is a finite set of $\mathscr{L}_{\mathrm{T}}$-sentences, $\mathscr{F}$ is an approximating function and $\models^{\alpha}_{\mathrm{T}} \Delta$ then $\models^{\alpha}_{\mathrm{T}} \mathscr{F}(\Delta)$.*

Proof. The proof is by induction on the length of the proof. It should be clear that if Δ is a template axiom then so is $\mathscr{F}(\Delta)$. It is also easy, but tedious, to check that the inference rules are not affected by $\mathscr{F}$. We will not do it here. $\square$

Remark 3.37. It should be clear that the lemma could be strengthen as to say that if Λ is a set of $\mathscr{L}_{\mathrm{T}}$-sentences such that if $\lambda \in \Lambda$ then $\mathscr{F}(\lambda) \in \Lambda$ for any approximating function $\mathscr{F}$, Δ a finite set of $\mathscr{L}_{\mathrm{T}}$-sentences, $\mathscr{F}$ an approximating function and $\Lambda \models^{\alpha}_{\mathrm{T}} \Delta$ then $\Lambda \models^{\alpha}_{\mathrm{T}} \mathscr{F}(\Delta)$.

Lemma 3.38. *If $\mathscr{F}_0$, $\mathscr{F}_1$, ... are approximating functions such that $|\mathscr{F}_i| \leqslant k$ and Γ a finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences then there is an approximating function $\mathscr{F}$ such that*

$$|\mathscr{F}| \leqslant \big(2^k - 1\big)|\Gamma|$$

and

$$\mathscr{F}(\Gamma) = \mathscr{F}(\mathscr{F}_i(\Gamma))$$

for all $i \in \omega$.

Proof. Let

$$\mathscr{F} = \mathscr{F}_{\delta_n} \circ \mathscr{F}_{\delta_{n-1}} \circ \ldots \circ \mathscr{F}_{\delta_1}$$

where $\delta_1, \ldots, \delta_n$ are all subformulas and subterms occurring in some formula in Γ at depth $\leqslant k$ in "the right order," i.e., if δ_i is a subformula or subterm of δ_j then $j < i$. Clearly $n \leqslant (2^k - 1)|\Gamma|$ and $\mathscr{F}(\Gamma) = \mathscr{F}(\mathscr{F}_i(\Gamma))$ for all i. $\square$

Definition 3.39. An approximating function

$$\mathscr{F} = \mathscr{F}_{\delta_n} \circ \ldots \circ \mathscr{F}_{\delta_1}$$

is said to be in *normal form* if $j < i$ for every i, j such that δ_i is a subformula or subterm of δ_j.

Observe that if $\mathscr{F}$ is an approximating function and $\mathscr{F}'$ is $\mathscr{F}_{\delta_k} \circ \ldots \circ \mathscr{F}_{\delta_0}$ where $\mathscr{F}_{\delta_0}, \ldots \mathscr{F}_{\delta_k}$ are the approximating functions in $\mathscr{F}$ ordered such that if δ_i is a subformula or subterm of δ_j then $i < j$, then $\mathscr{F}' \circ \mathscr{F}(\Gamma) = \mathscr{F}'(\Gamma)$ and $\mathscr{F}'$ is in normal form. Therefore if $\mathscr{F}(\Gamma)$ is provable so is $\mathscr{F}'(\Gamma)$.

From now on we will assume that all approximating functions are in normal form.

If $\mathscr{F}_1, \ldots, \mathscr{F}_k$ are approximating functions we can form *the uniform union*

$$\mathscr{F}_1 \star \ldots \star \mathscr{F}_k$$

of them which is any normal form of $\mathscr{F}_1 \circ \ldots \circ \mathscr{F}_k$. This definition is not unique, the reader may try to make it unique in a suitable way.

Please observe that if $\mathscr{F}_1$ and $\mathscr{F}_2$ are approximating functions and $\models_{\mathrm{T}} \mathscr{F}(\Gamma)$ then $\models_{\mathrm{T}} \mathscr{F}_1 \star \mathscr{F} \star \mathscr{F}_2(\Gamma)$ since $\mathscr{F}_1 \star \mathscr{F} \star \mathscr{F}_2(\mathscr{F}(\Gamma)) = \mathscr{F}_1 \star \mathscr{F} \star \mathscr{F}_2(\Gamma)$.

Lemma 3.40. *Let ψ be any ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence and $\mathscr{F}$ an approximating function (in normal form) such that $\mathscr{F}_\psi$ is in $\mathscr{F}$, then*

- *if ψ is $\neg\gamma$ then $\mathscr{F}(\neg\gamma) = \neg\mathscr{F}(\gamma)$,*
- *if ψ is $\gamma_1 \vee \gamma_2$ then $\mathscr{F}(\gamma_1 \vee \gamma_2) = \mathscr{F}(\gamma_1) \vee \mathscr{F}(\gamma_2)$ and*
- *if ψ is $\exists \mathsf{v}_i\, \gamma$ then $\mathscr{F}(\exists \mathsf{v}_i\, \gamma) = \exists \mathsf{v}_i\, \mathscr{F}(\gamma)$.*

Proof. The proof is more or less trivial and left to the reader. □

3.4.3 Semantics

We will now start to look at the semantics of template logic and prove a soundness theorem which implies the consistency of the logic. We end the section by also proving a completeness theorem.

Definition 3.41. An $\mathscr{L}_{\mathtt{T}}$-*structure* $\mathfrak{T}$ is a pair $\langle \mathfrak{T}_t, \mathfrak{T}_v \rangle$ of a set $\mathfrak{T}_t$ of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences and a map $\mathfrak{T}_v$ from the closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms into $\mathfrak{M}$.

Definition 3.42. If $\mathfrak{T}$ is an $\mathscr{L}_{\mathtt{T}}$-structure define $\mathtt{val}_{\mathfrak{T}}(t)$ for closed $\mathscr{L}_{\mathtt{T}}$-terms t inductively as follows:

$$\begin{aligned} \mathtt{val}_{\mathfrak{T}}(\widehat{r}) &=_{\mathrm{df}} \mathfrak{T}_v(r), \\ \mathtt{val}_{\mathfrak{T}}(\mathsf{S}(r)) &=_{\mathrm{df}} \mathsf{S}(\mathtt{val}_{\mathfrak{T}}(r)), \\ \mathtt{val}_{\mathfrak{T}}(r+s) &=_{\mathrm{df}} \mathtt{val}_{\mathfrak{T}}(r) + \mathtt{val}_{\mathfrak{T}}(s) \quad \text{and} \\ \mathtt{val}_{\mathfrak{T}}(r \cdot s) &=_{\mathrm{df}} \mathtt{val}_{\mathfrak{T}}(r) \cdot \mathtt{val}_{\mathfrak{T}}(s). \end{aligned}$$

Definition 3.43. If $\mathfrak{T}$ is an $\mathscr{L}_{\mathtt{T}}$-structure then define the predicate $\mathfrak{T} \vDash \varphi$ on $\mathscr{L}_{\mathtt{T}}$-sentences φ inductively as follows:

$$\begin{aligned} \mathfrak{T} \vDash t = r \quad &\text{iff} \quad \mathtt{val}_{\mathfrak{T}}(t) = \mathtt{val}_{\mathfrak{T}}(r), \\ \mathfrak{T} \vDash \widehat{\varphi} \quad &\text{iff} \quad \varphi \in \mathfrak{T}_t, \\ \mathfrak{T} \vDash \neg\gamma \quad &\text{iff} \quad \mathfrak{T} \nvDash \gamma, \\ \mathfrak{T} \vDash \gamma \vee \delta \quad &\text{iff} \quad \mathfrak{T} \vDash \gamma \text{ or } \mathfrak{T} \vDash \delta \quad \text{and} \\ \mathfrak{T} \vDash \exists \mathsf{v}_i \gamma \quad &\text{iff} \quad \text{there exists } a \in \mathfrak{M} \text{ such that } \mathfrak{T} \vDash \gamma[\mathsf{c}_a/\mathsf{v}_i]. \end{aligned}$$

We will now prove a soundness property for template logic. Let $\bigvee \Delta$ be

$$\delta_0 \vee (\delta_1 \vee (\ldots \vee \delta_k))$$

if $\Delta = \{ \delta_0, \ldots, \delta_k \}$ and $0 \neq 0$ if $\Delta = \emptyset$.

Proposition 3.44 (Soundness of template logic). *Let Δ be a finite set of $\mathscr{L}_{\mathtt{T}}$-sentences, let γ be $\bigvee \Delta$ and Λ any set of $\mathscr{L}_{\mathtt{T}}$-sentences. If $\mathfrak{T}$ is an $\mathscr{L}_{\mathtt{T}}$-structure making all the sentences of Λ true and $\Lambda \vDash_{\mathtt{T}} \Delta$ then $\mathfrak{T} \vDash \gamma$.*

Proof. We have to check that all the axioms of template logic are true in all $\mathscr{L}_\mathrm{T}$-structures and that all the inference rules are sound, i.e., if the premises of a rule are true in some $\mathscr{L}_\mathrm{T}$-structure then the conclusion is also true in the same $\mathscr{L}_\mathrm{T}$-structure. The axioms are quite obvious true and the inference rules are also easy to check; we only prove that $\mathfrak{M}$-rule$_t$ is sound. Suppose $\mathfrak{T} \models \neg\psi[\mathsf{c}_a/\mathsf{v}_i]$ for all $a \in \mathfrak{M}$ and $\mathfrak{T} \models \exists \mathsf{v}_i\, \psi$, then there is an $a \in \mathfrak{M}$ such that $\mathfrak{T} \models \psi[\mathsf{c}_a/\mathsf{v}_i]$ which is a contradiction. □

Definition 3.45. An $\mathscr{L}_\mathrm{T}$-sentence γ is said to be *true* in $\mathfrak{M}$ if $\mathfrak{T} \models \gamma$ for all $\mathscr{L}_\mathrm{T}$-structures $\mathfrak{T}$.

Corollary 3.46. *If* $\models_{\mathrm{T}} \Gamma$ *then* $\bigvee \Gamma$ *is true in* $\mathfrak{M}$.

Now, we easily get the consistency of template logic.

Corollary 3.47. *In any model* $\mathfrak{M}$ *we have* $\not\models_{\mathrm{T}} 0 \neq 0$.

Proof. Use Corollary 3.46 and the fact that for any template structure $\mathfrak{T}$ we have $\mathfrak{T} \not\models 0 \neq 0$. □

Proposition 3.48 (Completeness of template logic). *If* ψ *is an* $\mathscr{L}_\mathrm{T}$*-sentence true in* $\mathfrak{M}$ *then* $\models_{\mathrm{T}} \psi$.

Proof. Suppose $\not\models_{\mathrm{T}} \psi$ then $\{\, \neg\psi \,\}$ is a consistent set in template logic and could be extended to a maximally consistent set Λ in the same way as when we constructed satisfaction classes from consistent sets in $\mathfrak{M}$-logic.

Define the $\mathscr{L}_\mathrm{T}$-structure $\mathfrak{T}$ by letting

$$\mathfrak{T}_t = \{\, \varphi \mid \varphi \text{ is a } {}^*\!\mathscr{L}_{\mathfrak{M}}\text{-sentence and } \widehat{\varphi} \in \Lambda \,\}$$

and

$$\mathfrak{T}_v(t) = a \quad \text{iff} \quad \widehat{t} = \mathsf{c}_a \in \Lambda.$$

We prove that

$$\varphi \in \Lambda \quad \text{iff} \quad \mathfrak{T} \models \varphi$$

by induction on the construction of φ. To handle the case when φ is atomic we first prove that if t is a closed $\mathscr{L}_\mathrm{T}$-term then $t = \mathsf{c}_a \in \Lambda$ iff $\mathtt{val}_{\mathfrak{T}}(t) = a$.

Suppose t is $\widehat{r}$, then the claim is trivially true from the definition of $\mathfrak{T}_v$. If t is $\mathsf{S}(r)$ and $t = \mathsf{c}_a \in \Lambda$, by Axiom12$_t$, there is a $b \in \mathfrak{M}$ such that $r = \mathsf{c}_b \in \Lambda$ and $\mathsf{S}^{\mathfrak{M}}(b) = a$. By the induction hypothesis $\mathtt{val}_{\mathfrak{T}}(r) = b$ so $\mathtt{val}_{\mathfrak{T}}(t) = \mathsf{S}^{\mathfrak{M}}(a) = b$. The case when the term is $r + s$ or $r \cdot s$ is handled in a similar way.

If $t = r \in \Lambda$ then by Axiom12 (and Axiom3, Axiom4 and Axiom5) there is an $a \in \mathfrak{M}$ such that $t = \mathsf{c}_a \in \Lambda$ and $r = \mathsf{c}_a \in \Lambda$. By the fact proved above $\mathtt{val}_{\mathfrak{T}}(t) = \mathtt{val}_{\mathfrak{T}}(r)$, so $\mathfrak{T} \models t = r$.

On the other hand if $\mathfrak{T} \models t = r$ then $\mathtt{val}_{\mathfrak{T}}(t) = \mathtt{val}_{\mathfrak{T}}(r) = a$ for some $a \in \mathfrak{M}$. Since, by the maximality of Λ, there are b, $d \in \mathfrak{M}$ such that $t = \mathsf{c}_b \in \Lambda$ and $r = \mathsf{c}_d \in \Lambda$ we have, by the fact proved above, $b = d = a$. Therefore $t = r \in \Lambda$.

If $\varphi \vee \psi \in \Lambda$ then either $\varphi \in \Lambda$ or $\psi \in \Lambda$ by the maximality of Λ and so, by the induction hypothesis, either $\mathfrak{T} \models \varphi$ or $\mathfrak{T} \models \psi$, either way $\mathfrak{T} \models \varphi \vee \psi$. On the other hand, if $\varphi \vee \psi \notin \Lambda$ then neither φ nor ψ is in Λ so by the induction hypothesis $\mathfrak{T} \nvDash \varphi$ and $\mathfrak{T} \nvDash \psi$ which implies that $\mathfrak{T} \nvDash \varphi \vee \psi$.

If $\neg\varphi \in \Lambda$ then $\varphi \notin \Lambda$ by the consistency of Λ and by the induction hypothesis $\mathfrak{T} \nvDash \varphi$, therefore $\mathfrak{T} \vDash \neg\varphi$. If $\neg\varphi \notin \Lambda$ then $\varphi \in \Lambda$ so $\mathfrak{T} \vDash \varphi$ and $\mathfrak{T} \nvDash \neg\varphi$.

If $\exists \mathsf{v}_i\, \varphi \in \Lambda$ then there exists $a \in \mathfrak{M}$ such that $\varphi[\mathsf{c}_a/\mathsf{v}_i] \in \Lambda$ and so $\mathfrak{T} \vDash \varphi[\mathsf{c}_a/\mathsf{v}_i]$ and $\mathfrak{T} \vDash \exists \mathsf{v}_i\, \varphi$. And if $\exists \mathsf{v}_i\, \varphi \notin \Lambda$ then $\varphi[\mathsf{c}_a/\mathsf{v}_i] \notin \Lambda$ for all $a \in \mathfrak{M}$. By the induction hypothesis $\mathfrak{T} \nvDash \varphi[\mathsf{c}_a/\mathsf{v}_i]$ for all $a \in \mathfrak{M}$ so $\mathfrak{T} \nvDash \exists \mathsf{v}_i\, \varphi$. □

To sum up this section, the main result is that for an $\mathscr{L}_{\mathrm{T}}$-sentence γ we have that γ is true in $\mathfrak{M}$ iff $\vdash_{\mathrm{T}} \gamma$.

3.4.4 A link between $\mathfrak{M}$-logic and template logic

In this section we prove that finite provability in $\mathfrak{M}$-logic implies provability of some approximation in template logic. Since template logic is consistent (Theorem 3.44) this will imply that $\mathfrak{M}$-logic is consistent.

Proposition 3.49. *There is a (recursive) function $G : \omega \to \omega$ such that if $\vdash^{n}_{\mathfrak{M}} \Gamma$ then there is an approximating function $\mathscr{F}$ such that $|\mathscr{F}| \leqslant G(n)$ and $\vdash_{\mathrm{T}} \mathscr{F}(\Gamma)$.*

Proof. If $n = 1$ then Γ is an axiom. It is easy to see that an approximating function of length 9 is enough to make $\mathscr{F}(\Gamma)$ into an axiom of template logic.

Suppose we have defined G for all values $\leqslant n$ and take a proof of height $n+1$, by the induction hypothesis we get a proof of some approximation of the premises of the last inference of the proof, with the approximating functions of length $\leqslant G(n)$. Suppose the last inference is

Weak:

$$\frac{\Lambda}{\Lambda, \varphi}$$

By the induction hypothesis we have an approximating function $\mathscr{F}$ such that $\vdash_{\mathrm{T}} \mathscr{F}(\Lambda)$, by Weak$_t$ we get a proof of $\mathscr{F}(\Lambda), \mathscr{F}(\varphi)$ which is the same as $\mathscr{F}(\Lambda, \varphi)$, therefore $G(n+1) = G(n)$ is enough for this case.

$\vee$**I1 (and** $\vee$**I2):**

$$\frac{\Lambda, \varphi}{\Lambda, \varphi \vee \psi}$$

By the induction hypothesis we get an approximating function $\mathscr{F}_0$ such that

$$\vdash_{\mathrm{T}} \mathscr{F}_0(\Lambda, \varphi).$$

Let $\mathscr{F} = \mathscr{F}_0 \star \mathscr{F}_{\varphi\vee\psi}$ then $\mathscr{F}(\Lambda, \varphi)$ is provable and by $\vee$I1$_t$

$$\mathscr{F}(\Lambda), \mathscr{F}(\varphi) \vee \mathscr{F}(\psi) = \mathscr{F}(\Lambda, \varphi \vee \psi)$$

is provable. Therefore $G(n+1) = G(n) + 1$ is enough for this case.

$\vee$**I3:**

$$\frac{\Lambda, \neg\varphi \quad \Lambda, \neg\psi}{\Lambda, \neg(\varphi \vee \psi)}$$

By the induction hypothesis we get approximating functions $\mathscr{F}_0$ and $\mathscr{F}_1$ such that $\mathscr{F}_0(\Lambda, \neg\varphi)$ and $\mathscr{F}_1(\Lambda, \neg\psi)$ are both provable. Let

$$\mathscr{F} = \mathscr{F}_0 \star \mathscr{F}_1 \star \mathscr{F}_{\varphi \vee \psi} \star \mathscr{F}_{\neg(\varphi \vee \psi)} \star \mathscr{F}_{\neg\varphi} \star \mathscr{F}_{\neg\psi},$$

by Lemma 3.36 $\mathscr{F}(\Lambda, \neg\varphi)$ and $\mathscr{F}(\Lambda, \neg\psi)$ are both provable. Since

$$\mathscr{F}(\Lambda, \neg\varphi) = \mathscr{F}(\Lambda), \neg\mathscr{F}(\varphi) \text{ and } \mathscr{F}(\Lambda, \neg\psi) = \mathscr{F}(\Lambda), \neg\mathscr{F}(\psi)$$

we have by $\vee \mathrm{I3}_t$ a proof of

$$\mathscr{F}(\Lambda), \neg(\mathscr{F}(\varphi) \vee \mathscr{F}(\psi)) = \mathscr{F}(\Lambda, \neg(\varphi \vee \psi)).$$

Therefore $G(n+1) = 2G(n) + 4$ is enough for this case.

¬I:

$$\frac{\Lambda, \varphi}{\Lambda, \neg\neg\varphi}$$

By the induction hypothesis we get an approximating function $\mathscr{F}_0$ such that

$$\mathscr{F}_0(\Lambda, \varphi)$$

is provable. Let

$$\mathscr{F} = \mathscr{F}_0 \star \mathscr{F}_{\neg\varphi} \star \mathscr{F}_{\neg\neg\varphi},$$

then

$$\mathscr{F}(\Lambda, \varphi) = \mathscr{F}(\Lambda), \mathscr{F}(\varphi)$$

is provable, so $\neg \mathrm{I}_t$ gives us a proof of

$$\mathscr{F}(\Lambda), \neg\neg\mathscr{F}(\varphi) = \mathscr{F}(\Lambda), \mathscr{F}(\neg\neg\varphi) = \mathscr{F}(\Lambda, \varphi).$$

Thus $G(n+1) = G(n) + 2$ is enough for this case.

Cut:

$$\frac{\Gamma, \varphi \quad \Gamma, \neg\varphi}{\Gamma}$$

By the induction hypothesis we get approximating functions $\mathscr{F}_0$ and $\mathscr{F}_1$ such that $\mathscr{F}_0(\Lambda, \varphi)$ and $\mathscr{F}_1(\Lambda, \neg\varphi)$ are provable. Let

$$\mathscr{F} = \mathscr{F}_0 \star \mathscr{F}_1 \star \mathscr{F}_{\neg\varphi},$$

then

$$\mathscr{F}(\Gamma, \varphi) = \mathscr{F}(\Gamma), \mathscr{F}(\varphi) \quad \text{and}$$
$$\mathscr{F}(\Gamma, \neg\varphi) = \mathscr{F}(\Gamma), \neg\mathscr{F}(\varphi)$$

are both provable. By Cut_t we get a proof of $\mathscr{F}(\Gamma)$. Thus $G(n+1) = 2G(n) + 1$ is enough for this case.

∃I:

$$\frac{\Lambda, \varphi[\mathsf{c}_a/\mathsf{v}_i]}{\Lambda, \exists \mathsf{v}_i\, \varphi}$$

By the induction hypothesis we get an approximating function $\mathscr{F}_0$ such that

$$\mathscr{F}_0(\Lambda, \varphi[\mathsf{c}_a/\mathsf{v}_i])$$

is provable. Let

$$\mathscr{F} = \mathscr{F}_0 \star \mathscr{F}_{\exists \mathsf{v}_i\, \varphi}$$

then

$$\mathscr{F}(\Lambda, \varphi[\mathsf{c}_a/\mathsf{v}_i]) = \mathscr{F}(\Lambda), \mathscr{F}(\varphi)[\mathsf{c}_a/\mathsf{v}_i]$$

is provable (the equality is Lemma 3.35) so by $\exists \mathrm{I}_t$ we get a proof of

$$\mathscr{F}(\Lambda), \exists \mathsf{v}_i\, \mathscr{F}(\varphi) = \mathscr{F}(\Lambda), \mathscr{F}(\exists \mathsf{v}_i\, \varphi) = \mathscr{F}(\Lambda, \exists \mathsf{v}_i\, \varphi).$$

Thus $G(n+1) = G(n) + 1$ is enough for this case.

$\mathfrak{M}$-rule:

$$\frac{\ldots\ \Lambda, \neg\varphi[\mathsf{c}_a/\mathsf{v}_i]\ \ldots\ {}_{a \in \mathfrak{M}}}{\Lambda, \neg\exists \mathsf{v}_i\, \varphi}$$

By the induction hypothesis we have approximating functions $\mathscr{F}_a$ such that

$$\mathscr{F}_a(\Lambda, \neg\varphi[\mathsf{c}_a/\mathsf{v}_i]) = \mathscr{F}_a(\Lambda), \mathscr{F}_a(\neg\varphi)[\mathsf{c}_a/\mathsf{v}_i]$$

are provable for all $a \in \mathfrak{M}$. Let $\mathscr{F}'$ be as in Lemma 3.38 and

$$\mathscr{F} = \mathscr{F}' \star \mathscr{F}_{\exists \mathsf{v}_i\, \varphi} \star \mathscr{F}_{\neg\exists \mathsf{v}_i\, \varphi},$$

then

$$\mathscr{F}(\Lambda), \mathscr{F}(\neg\varphi)[\mathsf{c}_a/\mathsf{v}_i] = \mathscr{F}(\Lambda), \neg\mathscr{F}(\varphi)[\mathsf{c}_a/\mathsf{v}_i]$$

are all provable by Lemma 3.36. By $\mathfrak{M}$-rule$_t$ we get a proof of

$$\mathscr{F}(\Lambda), \neg\exists \mathsf{v}_i\, \mathscr{F}(\varphi) = \mathscr{F}(\Lambda), \mathscr{F}(\neg\exists \mathsf{v}_i\, \varphi) = \mathscr{F}(\Lambda, \neg\exists \mathsf{v}_i\, \varphi).$$

Thus $G(n+1) = (n+2)(2^{G(n)} - 1) + 2$ is enough since if $\vdash^n_{\mathfrak{M}} \Gamma$ then $|\Gamma| \leqslant n+2$ and therefore $|\mathscr{F}'| \leqslant (n+2)(2^{G(n)} - 1)$ by Lemma 3.36.

Thus, if we define G recursively by

$$\begin{aligned} G(1) &= 9 \\ G(n+1) &= (n+2)(2^{G(n)} - 1) + 2 \quad \text{for } n \geqslant 1, \end{aligned}$$

then G satisfies the proposition. □

Definition 3.50. If Λ is a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences then $\mathrm{app}(\Lambda)$ is a set of $\mathscr{L}_{\mathrm{T}}$-sentences defined as

$$\mathrm{app}(\Lambda) = \{\, \mathscr{F}(\lambda) \mid \lambda \in \Lambda \text{ and } \mathscr{F} \text{ is an approximating function} \,\}.$$

Porism 3.51. *If Λ is a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences and $\Lambda \vdash^{\omega}_{\mathfrak{M}} \Gamma$ then $\mathrm{app}(\Lambda) \vdash_{\mathrm{T}} \mathscr{F}(\Gamma)$ for some approximating function $\mathscr{F}$.*

In fact we could strengthen the porism: Let Δ be any set of $\mathscr{L}_{\mathrm{T}}$-sentences closed under approximating functions, i.e., if $\psi \in \Delta$ and $\mathscr{F}$ is an approximating function then $\mathscr{F}(\psi) \in \Delta$. Let Λ be a set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences and k a natural numbers such that for every $\varphi \in \Lambda$ there exists an approximating function $\mathscr{F}$ such that $|\mathscr{F}| \leqslant k$ and $\mathscr{F}(\varphi) \in \Delta$. If $\Lambda \vdash^{\omega}_{\mathfrak{M}} \Gamma$ then there exists $\mathscr{F}$ such that $\Delta \vdash_{\mathrm{T}} \mathscr{F}(\Gamma)$.

3.4.5 The consistency of $\mathfrak{M}$-logic

Proposition 3.52. *If $\mathfrak{M}$ is recursively saturated and $\vdash_{\mathfrak{M}} \Gamma$ then there is an approximating function $\mathscr{F}$ such that $\vdash_{\mathrm{T}} \mathscr{F}(\Gamma)$.*

Proof. Combine Propositions 3.27 and 3.49. □

Proposition 3.53. *Suppose $\mathfrak{M}$ is recursively saturated, Λ is a definable set and there is a template structure $\mathfrak{T}$ such that*

$$\mathfrak{T} \vDash \lambda, \quad \textit{for all } \lambda \in \mathrm{app}(\Lambda)$$

then Λ is consistent in $\mathfrak{M}$-logic.

Proof. By Proposition 3.27 if $\Lambda \vdash_{\mathfrak{M}} \emptyset$ then $\Lambda \vdash^{\omega}_{\mathfrak{M}} \emptyset$. By Porism 3.51 there is a proof in template logic of $\emptyset$ from the set $\mathrm{app}(\Lambda)$, i.e., $\mathrm{app}(\Lambda) \vdash_{\mathrm{T}} \emptyset$. By Proposition 3.44 we then have $\mathfrak{T} \vDash 0 \neq 0$ which clearly is a contradiction. □

Corollary 3.54. *If $\mathfrak{M}$ is recursively saturated then $\mathfrak{M}$-logic is consistent, i.e., $\nvdash_{\mathfrak{M}} \emptyset$.*

In fact, we can prove something a bit stronger by carefully examine the proofs of Proposition 3.27, 3.49 and 3.44:

Theorem 3.55. *Let Λ be a set of ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentences such that the expanded structure $\langle \mathfrak{M}, \Lambda \rangle$ is recursively saturated, let Δ be a set of $\mathscr{L}_{\mathrm{T}}$-sentences closed under approximating functions and n a natural number such that if $\varphi \in \Lambda$ then there exists an approximating function $\mathscr{F}$ such that $|\mathscr{F}| \leqslant n$ and $\mathscr{F}(\varphi) \in \Delta$. Moreover let*

$$\Delta_k = \{\, \mathscr{F}(\varphi) \,|\, |\mathscr{F}| \leqslant k \wedge \varphi \in \Lambda \wedge \mathscr{F}(\varphi) \in \Delta \,\}.$$

If there exists template structures $\mathfrak{T}_k$ such that

$$\mathfrak{T}_k \vDash \delta \quad \textit{for all } \delta \in \Delta_k$$

then Λ is consistent in $\mathfrak{M}$-logic.

The next result is a, sort of, negative result. Usually it is expressed as $\mathfrak{M}$-logic admits full pathology.

Let δ_0 be $0 \neq 0$ and by induction define δ_{a+1} to be $\delta_a \vee \delta_a$ for all $a \in \mathfrak{M}$.

Proposition 3.56. *If $\mathfrak{M}$ is recursively saturated and countable and $a \in \mathfrak{M} \setminus \omega$ then $\mathfrak{M}$ admits a satisfaction class Σ such that $\delta_a \in \Sigma$.*

Proof. The approximations of δ_a are

$$\widehat{\delta_a},$$
$$\widehat{\delta_{a-1}} \vee \widehat{\delta_{a-1}},$$
$$(\widehat{\delta_{a-2}} \vee \widehat{\delta_{a-2}}) \vee (\widehat{\delta_{a-2}} \vee \widehat{\delta_{a-2}}),$$

and so on. The template structure $\mathfrak{T} = \langle \mathfrak{T}_v, \mathfrak{T}_t \rangle$ with $\mathfrak{T}_v(t) = 0$ for all closed terms t and $\mathfrak{T}_t = \{\, \delta_{a-k} : k \in \omega \,\}$ makes all these approximations true. Applying Proposition 3.53 gives us the proposition. □

Proposition 3.57. *If t is a term with no constants or multiplications at finite depth, e.g.,*

$$\mathsf{S}(\mathsf{S}(\ldots\mathsf{S}(0)\ldots))$$

with a nonstandard number of successor symbols, and $a \in \mathfrak{M} \setminus \omega$ then there exists a satisfaction class Σ such that $t = \mathsf{c}_a \in \Sigma$.

Proof. We have to prove that there is a template structure making all the approximations of $t = \mathsf{c}_a$ true. Define $\mathfrak{T}_v(t) = a$ and by induction define $\mathfrak{T}_v$ on all closed terms occurring in t at constant depth. If $\mathfrak{T}_v(r) = b$ and r is $\mathsf{S}(s)$ then define $\mathfrak{T}_v(s) = b - 1$ and if r is $s + s'$ let

$$\mathfrak{T}_v(s) = \mathfrak{T}_v(s') = b/2$$

if b is even and

$$\mathfrak{T}_v(s) = \frac{b+1}{2}, \quad \mathfrak{T}_v(s') = \frac{b-1}{2}$$

if b is odd. Let $\mathfrak{T}_v(\mathsf{c}_a) = a$ and for all closed terms r not occurring in t at finite depth let $\mathfrak{T}_v(r) = 0$. Finally, let $\mathfrak{T}_t = \{\, t = \mathsf{c}_a \,\}$.

It should now be clear that $\mathtt{val}_{\mathfrak{T}}(\mathscr{F}(t)) = a$, and so $\mathfrak{T} \vDash \mathscr{F}(t = \mathsf{c}_a)$ for any approximating function $\mathscr{F}$. Applying Proposition 3.53 gives us the result. $\square$

The proposition is false if we allow multiplication at finite depth in t, since if t is, for example,

$$\mathsf{S}(\mathsf{S}(r)) \cdot \mathsf{S}(\mathsf{S}(s))$$

and $a \in \mathfrak{M}$ is prime, i.e.,

$$\mathfrak{M} \vDash \mathsf{S}(0) < a \land \forall x, y\big(x \cdot y = a \to x = \mathsf{S}(0) \lor y = \mathsf{S}(0)\big),$$

then $\mathrel{|\!\!\approx}_{\mathfrak{M}} t \neq \mathsf{c}_a$.

The next result is a partial answer to Question 3.7.

Proposition 3.58. *Let $\sim$ be an $\mathscr{L}_{\mathfrak{M}}$-definable equivalence relation on $\mathtt{ClTerm}(\mathfrak{M})$ and*

$$E = \{\, t = r \mid t \sim r \,\}\,.$$

If $\mathfrak{M}_E$ is well-defined and the canonical map

$$f : \mathfrak{M} \to \mathfrak{M}_E, \quad a \mapsto \overline{\mathsf{c}_a}$$

is an isomorphism then there is a satisfaction class Σ such that

$$t \sim r \quad \textit{iff} \quad t = r \in \Sigma$$

for all closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms t and r.

Proof. We prove that the $\mathscr{L}_{\mathfrak{M}}$-definable set E is consistent in $\mathfrak{M}$-logic. We do this by defining a template structure $\mathfrak{T} = \langle \mathfrak{T}_t, \mathfrak{T}_v \rangle$ such that all $\mathscr{L}_{\mathrm{T}}$-sentences in $\mathrm{app}(E)$ are true in $\mathfrak{T}$.

Let

$$\mathfrak{T}_v(t) = f^{-1}(\bar{t}\,)$$

and $\mathfrak{T}_t = E$. We claim that

$$\mathfrak{T} \vDash \mathscr{F}(t = r)$$

for closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms t and r such that $t \sim r$ and any approximating function $\mathscr{F}$. To see this we first observe that for any closed ${}^*\mathscr{L}_{\mathfrak{M}}$-term t and any approximating function $\mathscr{F}$ we have

$$\mathfrak{T} \vDash \widehat{t} = \mathscr{F}(t).$$

This is proved by induction on $|\mathscr{F}|$. For $|\mathscr{F}| = 0$ it is trivial and for for the induction step all we have to do is to observe that

$$\begin{aligned}\mathtt{val}_{\mathfrak{T}}(\mathsf{S}(\widehat{t}\,)) = \mathsf{S}^{\mathfrak{M}}(\mathtt{val}_{\mathfrak{T}}(\widehat{t}\,)) = \mathsf{S}^{\mathfrak{M}}(f^{-1}(\bar{t}\,)) = f^{-1}(\mathsf{S}^{\mathfrak{M}_E}(\bar{t}\,)) = \\ f^{-1}(\overline{\mathsf{S}(t)}) = \mathtt{val}_{\mathfrak{T}}(\widehat{\mathsf{S}(t)})\end{aligned}$$

and similar for $+$ and $\cdot$. This means that when we substitute, for example, $\mathsf{S}(\widehat{t}\,)$ for $\widehat{\mathsf{S}(t)}$ the value of $\mathtt{val}_{\mathfrak{T}}$ does not change. Therefore

$$\mathtt{val}_{\mathfrak{T}}(\mathscr{F}(t)) = \mathtt{val}_{\mathfrak{T}}(\widehat{t}\,)$$

for any closed ${}^*\mathscr{L}_{\mathfrak{M}}$-term t and any approximating function $\mathscr{F}$. This proves the claim. □

3.4.6 Some auxiliary results

We will end this chapter by proving converse results of Proposition 3.53 and 3.49.

Proposition 3.59. *If Λ is a consistent set in $\mathfrak{M}$-logic of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences then there is a template structure $\mathfrak{T}$ making all sentences $\lambda \in \mathrm{app}(\Lambda)$ true.*

Proof. Since Λ is consistent it is contained in some satisfaction class Σ. Define $\mathfrak{T}$ to interpret the approximation symbols exactly as Σ sees them, i.e., define

$$\mathfrak{T}_t = \Sigma$$

and

$$\mathfrak{T}_v(\widehat{t}\,) = a \quad \text{iff} \quad t = \mathsf{c}_a \in \Sigma$$

It is clear that $\mathfrak{T} \vDash \widehat{\varphi}$ if $\varphi \in \Lambda$, therefore, all we have to prove is that

$$\mathfrak{T} \vDash \psi \quad \Rightarrow \quad \mathfrak{T} \vDash \mathscr{F}_\tau(\psi),$$

for any $\mathscr{L}_{\mathrm{T}}$-sentences ψ and any template symbols τ.

If τ is $\widehat{\neg\varphi}$ then by observing that

$$\mathfrak{T} \models \widehat{\neg\varphi} \quad \text{iff} \quad \mathfrak{T} \models \neg\widehat{\varphi}$$

we see that replacing all occurrences of $\widehat{\neg\varphi}$ by $\neg\widehat{\varphi}$ does not change the truth value of the sentence. In the same way we have

$$\begin{aligned} \mathfrak{T} \models \widehat{\varphi \vee \psi} \quad &\text{iff} \quad \mathfrak{T} \models \widehat{\varphi} \vee \widehat{\psi} \qquad \text{and} \\ \mathfrak{T} \models \widehat{\exists \mathsf{v}_i \varphi} \quad &\text{iff} \quad \mathfrak{T} \models \exists \mathsf{v}_i \widehat{\varphi}. \end{aligned}$$

Therefore, substituting $\widehat{\varphi} \vee \widehat{\psi}$ for $\widehat{\varphi \vee \psi}$ in a sentence does not change the truth value. The same holds for $\exists \mathsf{v}_i \widehat{\varphi}$ and $\widehat{\exists \mathsf{v}_i \varphi}$.

We also have that if τ is $\widehat{\mathsf{S}(t)}$ then

$$\mathfrak{T} \models \widehat{\mathsf{S}(t)} = \mathsf{S}(\widehat{t}\,)$$

so substituting $\mathsf{S}(\widehat{t}\,)$ for $\widehat{\mathsf{S}(t)}$does not change the truth value either. The same is true for addition and multiplication since

$$\begin{aligned} &\mathfrak{T} \models \widehat{t + r} = \widehat{t} + \widehat{r} \quad \text{and} \\ &\mathfrak{T} \models \widehat{t \cdot r} = \widehat{t} \cdot \widehat{r}. \end{aligned}$$

We have to check two more cases, the following observations will handle those:

$$\begin{aligned} &\mathfrak{T} \models \widehat{t = r} \quad \text{iff} \quad \mathfrak{T} \models \widehat{t} = \widehat{r} \quad \text{and} \\ &\mathfrak{T} \models \widehat{\mathsf{c}_a} = \mathsf{c}_a. \end{aligned}$$

□

We are now in a good position to prove a converse to Proposition 3.49.

Proposition 3.60. *If Γ is a finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, $\mathscr{F}$ is any approximating function and $\models_{\mathrm{T}} \mathscr{F}(\Gamma)$ then $\models_{\mathfrak{M}} \Gamma$.*

Proof. Suppose that $\not\models_{\mathfrak{M}} \Gamma$, then $\neg\Gamma$ is consistent in $\mathfrak{M}$-logic. By Proposition 3.59 there is a template structure $\mathfrak{T}$ satisfying $\mathrm{app}(\neg\Gamma)$, but if $\models_{\mathrm{T}} \mathscr{F}(\Gamma)$ then

$$\mathfrak{T} \models \bigvee \mathscr{F}(\Gamma),$$

by Proposition 3.44. Clearly, then there is a $\gamma \in \Gamma$ such that

$$\mathfrak{T} \models \mathscr{F}(\gamma),$$

therefore, $\mathfrak{T} \not\models \neg\mathscr{F}(\gamma)$ and so

$$\mathfrak{T} \not\models \mathscr{F} \star \mathscr{F}_{\neg\gamma}(\neg\gamma).$$

Since

$$\mathscr{F} \star \mathscr{F}_{\neg\gamma}(\neg\gamma) \in \mathrm{app}(\neg\Gamma)$$

this yields a contradiction. □

To sum up we illustrate the chapter by Figure 3.4.

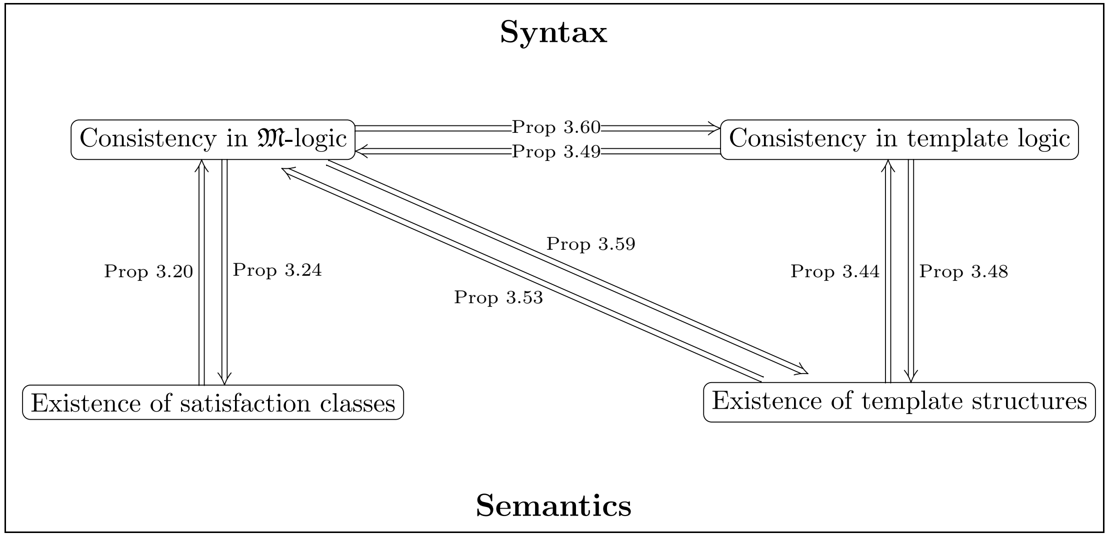

Figure 3.4: *A diagram showing the main results of Chapter 3.*

4

Weaker satisfaction classes

In this chapter we will study sets which fail to be satisfaction classes, but just merely; Axiom12 might be false in it, i.e., there might be terms t such that

$$\neg\exists \mathsf{v}_0(t = \mathsf{v}_0)$$

is in the set.

4.1 Free $\mathfrak{M}$-logic

We will try to answer the question:

> What happens if we remove Axiom12 from the axioms of $\mathfrak{M}$-logic?

Let us call a set of ${}^*\!\mathscr{L}_{\mathfrak{M}}$-sentences satisfying the alternative definition of satisfaction class given in Proposition 3.6 but with the word 'isomorphism' changed to 'isomorphic embedding' for a *free* satisfaction class. And $\mathfrak{M}$-logic without Axiom12 for *free* $\mathfrak{M}$-logic.[1]

Proposition 4.1. *Free $\mathfrak{M}$-logic corresponds to free satisfaction classes in the same way as $\mathfrak{M}$-logic corresponds to satisfaction classes, i.e., every maximally consistent set of sentences in free $\mathfrak{M}$-logic is a free satisfaction class and every free satisfaction class is a maximally consistent set of sentences.*

Proof. Assume X is a maximally consistent set in free $\mathfrak{M}$-logic. We have to prove that X satisfies the definition of a free satisfaction class, the only nontrivial parts are to prove that $\sim_X$ is an equivalence relation, that $\mathfrak{M}_X$ is well-defined and that the canonical map $f : \mathfrak{M} \to \mathfrak{M}_X$ is an isomorphic embedding.

Clearly, Axiom3, Axiom4 and Axiom5 implies that $\sim_X$ is an equivalence relation, furthermore Axiom6, Axiom7 and Axiom8 implies that $\mathfrak{M}_X$ is well-defined. That f is a homomorphism follows from Axiom9, Axiom10 and Axiom11. Finally, the injectivity follows from Axiom2.

[1] See Remark 4.5 for an explanation of the name.

For the converse, assume that Σ is a free satisfaction class. Σ is clearly maximally consistent if it is consistent since for every ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence φ either $\varphi \in \Sigma$ or $\neg\varphi \in \Sigma$. Thus, all we have to prove is that Σ is consistent in free $\mathfrak{M}$-logic, i.e., we have to check that if Δ is an axiom then $\bigvee \Delta \in \Sigma$ and that Σ is closed under all inference rules; the consistency then follows from the fact that there are ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences not in Σ.

From the fact that $\sim_\Sigma$ is an equivalence relation it is easy to see that the disjunctions of Axiom3, Axiom4 and Axiom5 all are in Σ. For Axiom6; assume that $t = r \in \Sigma$, by the well-definition of $\mathfrak{M}_\Sigma$

$$\overline{\mathsf{S}(t)} = \mathsf{S}^{\mathfrak{M}_\Sigma}(\,\overline{t}\,) = \mathsf{S}^{\mathfrak{M}_\Sigma}(\overline{r}) = \overline{\mathsf{S}(r)},$$

thus, $\mathsf{S}(t) = \mathsf{S}(r) \in \Sigma$. Similar for Axiom7 and Axiom8. Clearly, the disjunction of Axiom1 is in Σ, and if $a \neq b$ then

$$\overline{\mathsf{c}_a} = f(a) \neq f(b) = \overline{\mathsf{c}_b}$$

so $\mathsf{c}_a = \mathsf{c}_b \notin \Sigma$. Therefore, by the maximality of Σ we have $\mathsf{c}_a \neq \mathsf{c}_b \in \Sigma$.

Furthermore,

$$\overline{\mathsf{c}_{a+^{\mathfrak{M}}b}} = f(a +^{\mathfrak{M}} b) = f(a) +^{\mathfrak{M}_\Sigma} f(b) = \overline{\mathsf{c}_a} +^{\mathfrak{M}_\Sigma} \overline{\mathsf{c}_b} = \overline{\mathsf{c}_a + \mathsf{c}_b}.$$

Thus $\mathsf{c}_{a+^{\mathfrak{M}}b} = \mathsf{c}_a + \mathsf{c}_b \in \Sigma$ and similarly for Axiom9 and Axiom11.

That Σ is closed under the inference rules is proved by using properties (3.10), (3.11) and (3.12) of Σ, it is left to the reader. $\square$

A natural question now arises:

> Are there free satisfaction classes which are not satisfaction classes?

The answer is yes as we now will prove. We prove that if $a \in \mathfrak{M} \setminus \omega$ then we can find a free satisfaction class Σ such that $\neg\exists \mathsf{v}_0(n_a = \mathsf{v}_0) \in \Sigma$, where n_a is the closed term defined inductively as follows:

$$n_a = \begin{cases} 0 & \text{if } a = 0 \\ \mathsf{S}(n_b) & \text{if } a = \mathsf{S}^{\mathfrak{M}}(b). \end{cases}$$

The consistency criteria we worked out in Chapter 3 is too acute to handle this since in every $\mathscr{L}_{\mathsf{T}}$-structure $\mathfrak{T}$ we have

$$\mathfrak{T} \vDash \exists \mathsf{v}_0(\mathscr{F}(n_a) = \mathsf{v}_0)$$

for any approximating function $\mathscr{F}$. The solution to this problem is to redefine and make $\mathscr{L}_{\mathsf{T}}$-structures more general.

Definition 4.2. A *free* $\mathscr{L}_{\mathsf{T}}$-structure, $\mathfrak{T}$, is a pair, $\langle \mathfrak{T}_t, \mathfrak{T}_v \rangle$, of a set, $\mathfrak{T}_t$, of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences and a map, $\mathfrak{T}_v$, from the closed ${}^*\mathscr{L}_{\mathfrak{M}}$-terms into some $\mathscr{L}_A$-structure $\mathfrak{N} \supseteq \mathfrak{M}$.

Truth in a free $\mathscr{L}_{\mathsf{T}}$-structure is defined in the obvious way.

We state the consistency criteria in one of its simplest forms, but it should be evident that it could be strengthen as in Theorem 3.55.

Proposition 4.3. *Let φ be a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence and $\mathscr{F}_0$ an approximating function such that there exists a free $\mathscr{L}_{\mathsf{T}}$-structure making $\mathscr{F} \circ \mathscr{F}_0(\varphi)$ true for any approximating function $\mathscr{F}$. Then φ cannot be proved in free $\mathfrak{M}$-logic.*

Proof. Only some small modifications to the proofs of Proposition 3.44 and 3.49 is needed. The details are left to the reader. □

Let $\mathfrak{N} \supsetneq \mathfrak{M}$ and $b \in \mathfrak{N} \setminus \mathfrak{M}$. Define

$$\begin{aligned} \mathfrak{T}_v(n_{a-k}) &=_{\mathrm{df}} b - k \quad \text{for all } k \in \omega, \\ \mathfrak{T}_t &=_{\mathrm{df}} \emptyset, \\ \mathfrak{T} &=_{\mathrm{df}} \langle \mathfrak{T}_t, \mathfrak{T}_v \rangle \qquad \text{and} \\ \mathscr{F}_0 &=_{\mathrm{df}} \mathscr{F}_{\mathsf{v}_0} \circ \mathscr{F}_{n_a = \mathsf{v}_0} \circ \mathscr{F}_{\exists \mathsf{v}_0 (n_a = \mathsf{v}_0)} \circ \mathscr{F}_{\neg \exists \mathsf{v}_0 (n_a = \mathsf{v}_0)}. \end{aligned}$$

Clearly

$$\mathfrak{T} \models \mathscr{F} \circ \mathscr{F}_0(\neg \exists \mathsf{v}_0 (n_a = \mathsf{v}_0))$$

for every approximating function $\mathscr{F}$, since $\mathtt{val}_{\mathfrak{T}}(\mathscr{F}(n_a)) = b$ for every $\mathscr{F}$. Therefore, by the proposition, $\neg \exists \mathsf{v}_0 (n_a = \mathsf{v}_0)$ is consistent in free $\mathfrak{M}$-logic. By the usual construction we can find a maximally consistent set Σ including $\neg \exists \mathsf{v}_0 (n_a = \mathsf{v}_0)$ which by Proposition 4.1 is a free satisfaction class. Σ is not a satisfaction class since Axiom12 is not true.

Remark 4.4. Even though a free satisfaction class Σ is a satisfaction class if for every closed ${}^*\mathscr{L}_{\mathfrak{M}}$-term t there exists $a \in \mathfrak{M}$ such that $t = \mathsf{c}_a \in \Sigma$ and we get an equivalent definition of satisfaction class if we replace (3.9) in Definition 3.2 by

$$\exists x \ulcorner t = \mathsf{c}_x \urcorner \in \Sigma$$

for every closed ${}^*\mathscr{L}_{\mathfrak{M}}$-term t, we do *not* get an equivalent definition of a free satisfaction class by removing (3.9) in that definition. This is a consequence of the fact that the ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence

$$\mathsf{S}(\mathsf{c}_a) = \mathsf{S}(n_b) \wedge \mathsf{c}_a \neq n_b \tag{4.1}$$

is consistent in free $\mathfrak{M}$-logic for any $a, b \in \mathfrak{M}$ such that $b > \omega$, and therefore included in some free satisfaction class. But (3.6) implies that for any satisfaction class Σ in which 3.9 may fail the sentence (4.1) is false since if $\mathsf{S}(\mathsf{c}_a) = \mathsf{S}(n_b) \in \Sigma$ then $\mathsf{c}_{\mathsf{S}^{\mathfrak{M}}(a)} = \mathsf{S}(n_b) \in \Sigma$, thus (3.6) implies that $\mathsf{c}_a = n_b \in \Sigma$. To prove that the sentence (4.1) is consistent in free $\mathfrak{M}$-logic it is enough to construct an $\mathscr{L}_A$-structure $\mathfrak{N} \supsetneq \mathfrak{M}$ such that for some $d \in \mathfrak{N}$ we have $\mathsf{S}^{\mathfrak{N}}(d) = \mathsf{S}^{\mathfrak{M}}(a)$ and $d \neq a$. Then define an $\mathscr{L}_{\mathsf{T}}$-structure mapping n_{b-k} to $d - k$ for every $k \in \omega$.

Remark 4.5. We are using the term *free* since free $\mathfrak{M}$-logic is a sort of (positive) free logic, see [Lam01]. In fact the part of free $\mathfrak{M}$-logic where we only consider sentences of the form $\sigma[t_1, \ldots, t_k/\mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_k}]$, where σ is a $\mathscr{L}_{\mathfrak{M}}$-formula and $t_1, \ldots, t_k$ are closed ${}^*\!\mathscr{L}_{\mathfrak{M}}$-terms, is a (positive) free logic, with the existential predicate defined as

$$\mathrm{E}!\, t \leftrightarrow_{\mathrm{def}} \exists \mathsf{v}_0 (t = \mathsf{v}_0).$$

In fact, for any $\varphi(x)$ of this form and any closed ${}^*\!\mathscr{L}_{\mathfrak{M}}$-terms t and r we can prove (in this restricted free $\mathfrak{M}$-logic)

$$\varphi(t) \wedge t = r \to \varphi(r).$$

Therefore, it is easy to see that

$$\varphi(t) \wedge \mathrm{E}!\, t \to \exists x \varphi(x)$$

also is provable (in the same logic). See [Lam01] or [Ben99] for more information on free logics.

Question 4.6. *Which $\mathscr{L}_A$-structures $\mathfrak{N} \supseteq \mathfrak{M}$ are $\mathfrak{M}_\Sigma$ for some free satisfaction class Σ?*

Question 4.7. *Are there free satisfaction classes Σ, which are not satisfaction classes, cuh that the canonical map $f : \mathfrak{M} \to \mathfrak{M}_\Sigma$ is an elementary embedding?*

5

Stronger satisfaction classes

As we have seen in Chapter 3 some "pathological" examples arise in the study of satisfaction classes. For example, we can make the sentences δ_a and $n_a = \mathtt{c}_b$ true in a satisfaction class (if a, b are nonstandard).[1] The main question we will try to answer in this chapter is:

What do we need to remove such "pathological" examples?

To answer this question we have concentrated on extensions of $\mathfrak{M}$-logic. Any maximally consistent set of sentences in any of the extensions we will study in this chapter is a satisfaction class removing pathologies of a certain kind.

It should be remarked that this chapter is included to emphasise the vast amount of open questions in this area. There is a lot of work to be done, I have just scratched the surface. The big question of consistency of the extensions is a hard question. We know the answer to some of them but not to all, but a small remark is in order here:

Remark 5.1. Since every logic we will study have axioms and inference rules definable in $\mathtt{PA}$, consistency is a Σ^1_1 statement:

$$\exists X \exists \varphi \big(\mathtt{Sent}(\varphi) \wedge \varphi \notin X \wedge X \text{ includes all axioms} \\ \wedge X \text{ is closed under all inference rules} \big).$$

Therefore, in countable models, consistency of one of these extensions of $\mathfrak{M}$-logic could only depend on $\mathtt{Th}(\mathfrak{M})$ and not on other model theoretic properties of $\mathfrak{M}$, such as saturation properties, since if $\mathfrak{M}$-logic is consistent then $\mathfrak{M}$ is recursively saturated by Theorem 3.29 and so resplendent by Theorem 2.7.

Figure 5.1 is a summary of some of the pathologies and their "solutions." The sentences ϵ^φ_a are defined as follows:

$$\begin{array}{rcll} \epsilon^\varphi_0 & \text{is} & \neg(\varphi \vee \neg\varphi) & \quad\text{and} \\ \epsilon^\varphi_{a+1} & \text{is} & \epsilon_a \vee \epsilon_a. & \end{array}$$

[1] Remember that δ_0 is $0 \neq 0$, δ_{a+1} is $\delta_a \vee \delta_a$, n_0 is 0 and $n_{\mathsf{S}^{\mathfrak{M}}(a)}$ is $\mathsf{S}(n_a)$.

Pathology	Solution	Consistent?	Section
$n_a = \mathsf{c}_b$	Axioms: $\mathtt{Tr}_{\mathrm{At}}$	Yes	5.1
δ_a	Axioms: $\mathtt{Tr}_{\Delta_0}$	Yes	5.1
ϵ_a^{φ}	Rule: Prop	?	5.2
$\exists \mathsf{v}_0 \mathsf{v}_1 \dots \mathsf{v}_a \, 0 \neq 0$	Axioms: $\mathtt{Tr}_{\Sigma_1}$	Yes	5.1
$\left(\exists \mathsf{v}_0 \mathsf{v}_1 \dots \mathsf{v}_a \, \varphi\right) \leftrightarrow \neg\varphi$	Rules: $\mathrm{I}\exists^{\infty}$, $\mathfrak{M}^{\infty}$-rule	?	5.3
	Rule: Pred	?	5.5
$\left(\exists \mathsf{v}_0 \forall \mathsf{v}_1 \dots \exists \mathsf{v}_{2a} \, \varphi\right) \leftrightarrow \neg\varphi$	Rule: Skolem-rule	?	5.4
	Rule: Pred	?	5.5

Figure 5.1: *A summary of the pathological examples we will study in this chapter and their "solutions."*

In the table, a is nonstandard and φ is a sentence of high complexity (it is not Σ_k for any $k \in \omega$). The 'Solution' to a pathology tells us what we need to add to $\mathfrak{M}$-logic (either axioms of inference rules) to remove the pathology, i.e., to be able to prove the negation of the pathology. The 'Consistent?' column tells us if this logic is consistent or not; a 'Yes' means that in any recursively saturated model the logic is consistent and a '?' means that we do not know the answer. The column named 'Section' is a reference for where to read more about the solution, it is the section number in this chapter.

Question 5.2. *How should Figure 5.1 be completed?*

5.1 Partial Truth Definitions

In [Kay91] a satisfaction class is defined to extend the set

$$\mathtt{Tr}_{\mathrm{At}} = \{\, t = r \,|\, \mathfrak{M} \vDash \mathtt{val}(t) = \mathtt{val}(r) \,\}\,.$$

By Proposition 3.53 this set is consistent since we can define a template structure $\mathfrak{T} = \langle \mathfrak{T}_t, \mathfrak{T}_v \rangle$, with

$$\mathfrak{T}_t = \{\, t = r \,|\, \mathfrak{M} \vDash \mathtt{val}(t) = \mathtt{val}(r) \,\} \quad \text{and}$$
$$\mathfrak{T}_v = \mathtt{val},$$

where $\mathtt{val}$ is the valuation function definable in $\mathtt{PA}$. This is a template structure making all $\mathscr{L}_{\mathrm{T}}$-sentences in app($\mathtt{Tr}_{\mathrm{At}}$) true. In the same manner we can find template structures making all sentences in app($\mathtt{Tr}_{\Sigma_k}$) true.

Proposition 5.3. *The sets* $\mathtt{Tr}_{At}, \mathtt{Tr}_{\Delta_k}, \mathtt{Tr}_{\Sigma_k}$ *and* $\mathtt{Tr}_{\Pi_k}$ *are all consistent.*

Proof. It is clearly enough to prove that $\mathtt{Tr}_{\Sigma_k}$ is consistent for any $k \in \omega$. The sets are definable so, by Proposition 3.53, it suffices to find a template structure making the $\mathscr{L}_{\mathrm{T}}$-sentences in app($\mathtt{Tr}_{\Sigma_k}$) true.

Let $\mathfrak{T} = \langle \mathfrak{T}_v, \mathfrak{T}_t \rangle$ be such that $\mathfrak{T}_v(t) = a$ iff $\mathfrak{M} \vDash \mathtt{val}(t) = \mathsf{c}_a$ and let $\mathfrak{T}_t = \mathtt{Tr}_{\Sigma_k}$. By an easy induction it is easy to see that this structure satisfies the condition. The induction is left to the reader, but we remark that the properties of $\mathtt{val}$ and $\mathtt{Tr}_{\Sigma_k}$ in Section 2.4 are used heavily. □

5.2 Closure under propositional logic

Satisfaction classes closed under nonstandard propositional proofs, in the sense that if $\mathfrak{M}$ thinks φ is provable in propositional logic from sentences in Σ then $\varphi \in \Sigma$, is the next object of study. Firstly, we have to define what it means for a model to think something is provable in propositional logic, i.e., we need some formula expressing propositional provability.

Definition 5.4. If $\sigma(x)$ is any formula then $\mathtt{PropPrf}_\sigma(y)$ is defined to be the formula

$$\exists x\Big([x]_{\mathrm{len}(x)-1} = y \wedge \forall i{<}\,\mathrm{len}(x)\big[\mathtt{Sent}([x]_i) \wedge \big[\mathtt{Ax}([x]_i) \vee \sigma([x]_i) \\ \vee \exists j,k{<}i\exists z\big([x]_j = \ulcorner z \to [x]_i\urcorner \wedge [x]_k = z\big)\big]\big]\Big),$$

where $\mathtt{Ax}(x)$ is a formula defining the axioms of propositional logic.[2]

The formula $\mathtt{PropPrf}_\sigma(y)$ says that there exists a sequence of sentences such that every element in the sequence is a ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence φ and either an axiom, satisfying σ, or a result of applying Modus Ponens to other sentences occurring in the sequence prior to φ.

This section is about satisfaction classes closed under this relation, in the sense that

$$\mathfrak{M} \vDash \forall x\big(\mathtt{PropPrf}_{x\in\Sigma}(x) \to x \in \Sigma\big).$$

For simplicity we will write $\Lambda \mathrel{|\!\!\!\approx^{*}_{\mathrm{p}}} \varphi$ to mean $\mathfrak{M} \vDash \mathtt{PropPrf}_{x\in\Lambda}(\varphi)$, but please do remember that all propositional proofs are "inside" the model $\mathfrak{M}$.

It is important to observe that we have a sort of compactness theorem even for nonstandard proofs:

Proposition 5.5. *Let $\mathfrak{M}^+$ be any expansion of $\mathfrak{M}$, $\sigma(x)$ a formula in the language of $\mathfrak{M}^+$ and $\sigma \mathrel{|\!\!\!\approx^{*}_{\mathrm{p}}} \varphi$, then there exists an $\mathscr{L}_{\mathfrak{M}}$-formula $\delta(x)$ such that*

$$\mathfrak{M}^+ \vDash \forall x\big(\delta(x) \to \sigma(x)\big) \wedge \exists x\forall y\big(\delta(y) \to y < x\big) \wedge \mathtt{PropPrf}_\delta(\varphi).$$

Proof. Let p be a proof of φ from σ, i.e., a witness for the existential sentence $\mathtt{PropPrf}_\sigma(\varphi)$. Define $\delta(x)$ to be

$$\exists i{<}\,\mathrm{len}(p)\Big([p]_i = x \wedge \neg\,\mathtt{Ax}(x) \wedge \forall j,k{<}\,\mathrm{len}(p)\forall y\big([p]_j = \ulcorner y \to x\urcorner \to [p]_k \neq y\big)\Big),$$

saying that x is a sentence in the proof p but it is not a result of Modus Ponens, neither an axiom of propositional logic. Clearly, this $\mathscr{L}_{\mathfrak{M}}$-formula has the desired property. □

[2] The axioms could be chosen in a variety of ways. Use your favourite axiomatisation.

Define $\bigvee\Gamma$, for any finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences Γ, in the following way: let $\bigvee\Gamma$ be the ${}^*\mathscr{L}_{\mathfrak{M}}$-sentence

$$\gamma_0 \vee \big(\gamma_2 \vee \big(\gamma_3 \vee \big(\dots \gamma_{k-1}\big)\big)\big)$$

where $\gamma_i < \gamma_{i+1}$ for all $i < k-1$ and $\gamma_0, \dots, \gamma_{k-1}$ enumerates the set Γ.

We will define a new logic, extending $\mathfrak{M}$-logic, that corresponds to satisfaction classes closed under propositional proofs. We call it $\mathfrak{M}_{\mathrm{p}}$-logic and it is similar to $\mathfrak{M}$-logic, with the important difference that Weak, $\vee$I1, $\vee$I2, $\vee$I3, $\neg$I and Cut are replaced by the single rule:

$$\frac{\dots\ \Lambda, \varphi_i\ \dots\ {}_{i<a}}{\Gamma} \qquad \text{if } \Big\{ \bigvee\Lambda \vee \varphi_0, \dots, \bigvee\Lambda \vee \varphi_{a-1} \Big\} \vdash^*_{\mathrm{p}} \bigvee\Gamma, \tag{Prop}$$

where $a \in \mathfrak{M}$ and φ_i, for $i < a$, are ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences.

This means that the inference rules of $\mathfrak{M}_{\mathrm{p}}$-logic are Prop, $\exists$I and $\mathfrak{M}$-rule. Please observe that we still restrict the sets of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences we derive to be actually finite.

It is easy to see that Weak, $\vee$I1, $\vee$I2, $\vee$I3, $\neg$I and Cut are all derivable from the single rule Prop, therefore, $\mathfrak{M}_{\mathrm{p}}$-logic is an extension of $\mathfrak{M}$-logic.

Let $\vdash_{\mathfrak{M}_{\mathrm{p}}}$ denote provability in $\mathfrak{M}_{\mathrm{p}}$-logic and let $\vdash^{\alpha}_{\mathfrak{M}_{\mathrm{p}}}$ mean provability with a proof of height less than α, analogous to the $\vdash_{\mathfrak{M}}$ and $\vdash^{\alpha}_{\mathfrak{M}}$ relations.

We will show that $\mathfrak{M}_{\mathrm{p}}$-logic is actually "finite" in the sense that if something is provable then it is provable by a finite height proof (as $\mathfrak{M}$-logic also is). This will be true for all logics studied in this chapter, but since all proofs follow the same line, we will only prove it for $\mathfrak{M}_{\mathrm{p}}$-logic.

We define formulas $\mathtt{Pf}'_k(x)$ as follows. Let $\mathtt{Pf}'_1(x)$ be the formula defining the axioms of $\mathfrak{M}_{\mathrm{p}}$-logic, i.e., $\mathfrak{M} \models \mathtt{Pf}'_1(a)$ iff a is $\bigvee\Gamma$ for Γ an axiom of $\mathfrak{M}_{\mathrm{p}}$-logic. Furthermore, let

$$\begin{aligned} \mathtt{Pf}'_{k+1}(x) = {} & \mathtt{PropPrf}_{\mathtt{Pf}'_k}(x) \\ & \vee \exists y, z, i\big(\mathtt{Sent}(y) \wedge \mathtt{Sent}(\ulcorner \exists \mathsf{v}_i\, z \urcorner) \wedge x = \ulcorner \exists \mathsf{v}_i\, z \vee y \urcorner \\ & \quad \wedge \exists y\, \mathtt{Pf}'_k(\ulcorner z[\mathsf{c}_y/\mathsf{v}_i] \vee y \urcorner)\big) \\ & \vee \exists y, z, i\big(\mathtt{Sent}(y) \wedge \mathtt{Sent}(\ulcorner \exists \mathsf{v}_i\, z \urcorner) \wedge x = \ulcorner \neg\exists \mathsf{v}_i\, z \vee y \urcorner \\ & \quad \wedge \forall y\ \mathtt{Pf}'_k(\ulcorner \neg z[\mathsf{c}_y/\mathsf{v}_i] \vee y \urcorner)\big), \end{aligned}$$

for $k > 1$.

These formulas "code" proofs of finite height in $\mathfrak{M}_{\mathrm{p}}$-logic in the following way:

Lemma 5.6. *If Γ is a finite set of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences, φ is $\bigvee\Gamma$ and $k > 0$, then*

$$\begin{aligned} \vdash^{k}_{\mathfrak{M}_{\mathrm{p}}} \Gamma \quad &\Longrightarrow \quad \mathfrak{M} \models \mathtt{Pf}'_k(\varphi) \qquad \textit{and} \\ \mathfrak{M} \models \mathtt{Pf}'_k(\varphi) \quad &\Longrightarrow \quad \vdash^{3k-2}_{\mathfrak{M}_{\mathrm{p}}} \varphi. \end{aligned}$$

Proof. The proof is by induction on k. For the base case, assume that $\vdash^1_{\mathfrak{M}_\mathrm{p}} \Gamma$ then Γ is an axiom of $\mathfrak{M}_\mathrm{p}$-logic and $\mathfrak{M} \models \mathtt{Pf}'_1(\varphi)$. On the other hand, if $\mathfrak{M} \models \mathtt{Pf}'_1(\varphi)$ then, let Λ be the axiom such that $\bigvee \Lambda$ is φ, since

$$\bigvee \Lambda \vdash^*_\mathrm{p} \varphi,$$

we have $\vdash^2_{\mathfrak{M}_\mathrm{p}} \varphi$.

For the induction step assume that the lemma holds for $k \leqslant n$. We will only prove the Prop cases, the $\exists$I and the $\mathfrak{M}$-rule cases are left to the reader.

Assume that $\vdash^{n+1}_{\mathfrak{M}_\mathrm{p}} \Gamma$ and that the last inference in the proof of Γ is Prop, then there is a finite set Λ and ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences $\varphi_0, \dots, \varphi_{a-1}$ such that

$$\left\{ \bigvee \Lambda \vee \varphi_0, \dots, \bigvee \Lambda \vee \varphi_{a-1} \right\} \vdash^*_\mathrm{p} \bigvee \Gamma,$$

and $\vdash^n_{\mathfrak{M}_\mathrm{p}} \Lambda, \varphi_i$ for all $i < a$. By the induction hypothesis,

$$\mathfrak{M} \models \mathtt{Pf}'_n(\bigvee \Lambda, \varphi_i),$$

so, clearly,

$$\mathtt{Pf}'_n \vdash^*_\mathrm{p} \bigvee \Lambda \vee \varphi_i,$$

for all $i < a$. Thus, $\mathtt{Pf}'_n \vdash^*_\mathrm{p} \varphi$, and so $\mathfrak{M} \models \mathtt{Pf}'_{n+1}(\varphi)$.

On the other hand, if $\mathfrak{M} \models \mathtt{Pf}'_{n+1}(\varphi)$ and

$$\mathfrak{M} \models \mathtt{Pf}'_n \vdash^*_\mathrm{p} \varphi,$$

then there exists, by Proposition 5.5, ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences $\varphi_0, \dots, \varphi_{b-1}$ such that

$$\mathfrak{M} \models \mathtt{Pf}'_n(\varphi_i)$$

for all $i < b$ and

$$\{ \varphi_i \}_{i<b} \vdash^*_\mathrm{p} \varphi.$$

Therefore, by the induction hypothesis,

$$\vdash^{3n-2}_{\mathfrak{M}_\mathrm{p}} \varphi_i$$

for all $i < b$, and so, by using Prop,

$$\vdash^{3n-1}_{\mathfrak{M}_\mathrm{p}} \varphi. \qquad \square$$

The lemma tells us that

$$\vdash^\omega_{\mathfrak{M}_\mathrm{p}} \Gamma \quad \text{iff} \quad \text{there exists } k \in \omega \text{ such that } \mathfrak{M} \models \mathtt{Pf}'_k(\bigvee \Gamma),$$

for any finite set Γ of ${}^*\mathscr{L}_{\mathfrak{M}}$-sentences.

Lemma 5.7. *If $\mathfrak{M}$ is recursively saturated and $\vdash_{\mathfrak{M}_\mathrm{p}} \Gamma$ then $\vdash^\omega_{\mathfrak{M}_\mathrm{p}} \Gamma$.*

Proof. Assume the lemma is false and let Γ be such that

$$\vdash^{\omega+1}_{\mathfrak{M}_{\mathrm{p}}} \Gamma \quad \text{but} \quad \nvdash^{\omega}_{\mathfrak{M}_{\mathrm{p}}} \Gamma,$$

we can find such Γ as in Lemma 3.27. Either the last rule in the proof of height ω of Γ is $\mathfrak{M}$-rule or Prop. For the first case the last inference is

$$\frac{\ldots \ \Delta, \neg\psi[\mathsf{c}_a/\mathsf{v}_i] \ \ldots \ {}_{a\in\mathfrak{M}}}{\Delta, \neg\exists\mathsf{v}_i\,\psi}.$$

Define the type

$$p(x) = \left\{ \neg\,\mathtt{Pf}'_k(\ulcorner\neg\psi[\mathsf{c}_x/\mathsf{v}_i] \vee \varphi\urcorner) \,\middle|\, k \in \omega \right\},$$

where φ is $\bigvee \Delta$.

For the second case the last inference is

$$\frac{\ldots \ \Lambda, \varphi_i \ \ldots \ {}_{i<a}}{\Gamma},$$

where

$$\{\, \psi \vee \varphi_0, \ldots, \psi \vee \varphi_{a-1} \,\} \vdash^{*}_{\mathrm{p}} \varphi,$$

ψ is $\bigvee \Lambda$, φ is $\bigvee \Gamma$ and a is nonstandard. By Proposition 5.5 there is a $\mathscr{L}_{\mathfrak{M}}$-definable subset Δ of $\{\, \psi \vee \varphi_i \,\}_{i<a}$ such that

$$\Delta \vdash^{*}_{\mathrm{p}} \varphi$$

Let $b \in \mathfrak{M}$ enumerate Δ such that

$$\{\, [b]_i \,\}_{i<\operatorname{len} b} = \Delta$$

and define the type

$$p(x) = \left\{ \neg\,\mathtt{Pf}'_k(\ulcorner\psi \vee (b)_x\urcorner) \,\middle|\, k \in \omega \right\} \cup \{\, x < \operatorname{len} b \,\}.$$

In either case $p(x)$ is a non-realized recursive type contradicting the recursive saturation of $\mathfrak{M}$. $\square$

The Lemma tells us that to prove the consistency of $\mathfrak{M}_{\mathrm{p}}$-logic we only need to prove that

$$\forall k{\in}\omega\, \mathfrak{M} \models \neg\,\mathtt{Pf}'_k(\ulcorner 0 \neq 0\urcorner).$$

We have, however, not succeeding in doing so.[3]

[3] In a preliminary draft of this thesis a proof of the consistency was presented, but it turned out to be erroneous.

Remark 5.8. It is easy to see that there are models $\mathfrak{M}$ where $\mathfrak{M}_{\mathrm{p}}$-logic is consistent. Let $\mathbb{N}$ be the standard model of PA and let Σ_0 be the standard satisfaction class on $\mathbb{N}$, i.e.,

$$\Sigma_0 = \{ \ulcorner\varphi\urcorner \mid \varphi \in \mathrm{ElDiag}(\mathbb{N}) \} .$$

Clearly, Σ_0 is closed under propositional proofs; thus,

$$\langle \mathbb{N}, \Sigma_0 \rangle \models \mathtt{SatCl}(\Sigma_0) \wedge \forall x \big(\mathtt{PropPrf}_{x \in \Sigma_0}(x) \to x \in \Sigma_0 \big).$$

Therefore, if

$$\langle \mathbb{N}, \Sigma_0 \rangle \prec \langle \mathfrak{M}, \Sigma \rangle$$

then Σ is a satisfaction class closed under propositional proofs.

Let us now instead prove that $\mathfrak{M}_{\mathrm{p}}$-logic corresponds to satisfaction classes closed under propositional proofs.

Proposition 5.9. *Satisfaction classes closed under propositional proofs are exactly the maximally consistent sets in* $\mathfrak{M}_{\mathrm{p}}$*-logic.*

Proof. Let Σ be a maximally consistent set. It is easy to see that it is a satisfaction class, just as we did in Chapter 3. We check that it is closed under propositional logic. Suppose $\Sigma \mathrel{|}\!\!\!{\frac{\ast}{\mathrm{p}}}\ \varphi$, then by the maximality either φ or $\neg\varphi$ is in Σ. If $\neg\varphi \in \Sigma$ then $\Sigma \vdash_{\mathfrak{M}_{\mathrm{p}}} \emptyset$, therefore, since Σ is consistent, $\varphi \in \Sigma$.

If Σ is a satisfaction class closed under propositional proofs then it is consistent in $\mathfrak{M}_{\mathrm{p}}$-logic, since it is closed under Prop, ∃I and $\mathfrak{M}$-rule. It is maximally consistent since for every φ either $\varphi \in \Sigma$ or $\neg\varphi \in \Sigma$. □

5.3 Infinite ∃I and $\mathfrak{M}$-rule

The inference rule Prop handles the propositional connectives in a satisfying way. In a first try to handle quantifiers we add two infinite versions of ∃I and $\mathfrak{M}$-rule:

$$\frac{\Gamma, \varphi[\mathsf{c}_{(a)_0}, \mathsf{c}_{(a)_1}, \ldots, \mathsf{c}_{(a)_b} / \mathsf{v}_{i_0}, \mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_b}]}{\Gamma, \exists \mathsf{v}_{i_0}, \mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_b}\, \varphi} \qquad (\mathrm{I}\exists^\infty)$$

and

$$\frac{\ldots\ \Gamma, \neg\varphi[\mathsf{c}_{(a)_0}, \mathsf{c}_{(a)_1}, \ldots, \mathsf{c}_{(a)_b} / \mathsf{v}_{i_0}, \mathsf{v}_{i_1}, \ldots, \mathsf{v}_{i_b}]\ \ldots\ {}_{a \in \mathfrak{M}}}{\Gamma, \neg\exists \mathsf{v}_{i_0}, \ldots, \mathsf{v}_{i_b}\, \varphi} \qquad (\mathfrak{M}^\infty\text{-rule})$$

where $b \in \mathfrak{M}$ may be nonstandard.

Definition 5.10 ([Kra76]). A satisfaction class Σ is ∃-*complete* if

$$\exists \mathsf{v}_{i_0}, \ldots, \mathsf{v}_{i_b}\, \varphi \in \Sigma \quad \text{iff} \quad \text{there exists } a \in \mathfrak{M} \text{ such that } \varphi[\mathsf{c}_{(a)_0}, \ldots, \mathsf{c}_{(a)_b} / \mathsf{v}_{i_0}, \ldots, \mathsf{v}_{i_b}] \in \Sigma.$$

The following proposition should now be easy to prove.

Proposition 5.11. *A set of* $^*\mathscr{L}_{\mathfrak{M}}$*-sentences is a maximally consistent set in* $\mathfrak{M}$*-logic with* $I\exists^\infty$ *and* $\mathfrak{M}^\infty$*-rule added iff it is a* ∃*-complete satisfaction class.*

Proof. Left to the reader. □

5.4 Skolem operators

It seems that we could extend the infinite quantifier rules even more. To be able to state this extended quantifier rule we need the notion of Skolem operators.

Let Q be a ${}^*\mathscr{L}_{\mathfrak{M}}$-definable sequence of quantifiers, i.e., $\mathsf{Q} \in \mathfrak{M}$ such that

$$\forall j < \operatorname{len}(\mathsf{Q}) \exists i \big([\mathsf{Q}]_j = \ulcorner \exists \mathsf{v}_i \urcorner \vee [\mathsf{Q}]_j = \ulcorner \forall \mathsf{v}_i \urcorner \big).$$

Define a function $f_{\mathsf{Q}} : \mathfrak{M} \mapsto \mathfrak{M}$ such that the $f_{\mathsf{Q}}(j)$th $\forall$-quantifier in Q is the first $\forall$-quantifier preceding (to the left of) the jth $\exists$-quantifier in Q. More formally; there is a $\mathtt{PA}$-definable function $F(x, y)$ such that the following is provable in $\mathtt{PA}$:

$$\begin{aligned} F(x, y) &= 0 \quad \text{if } x \text{ is not a sequence of quantifiers.} \\ F([], y) &= 0 \\ F(x, 0) &= 0 \\ F([\ulcorner \exists \mathsf{v}_i \urcorner] ^\frown x, \mathsf{S}(y)) &= F(x, y) \\ F([\ulcorner \forall \mathsf{v}_i \urcorner] ^\frown x, y) &= \mathsf{S}(F(x, y)) \quad \text{if } y \neq 0 \end{aligned}$$

Define $f_{\mathsf{Q}}(x) = F(\mathsf{Q}, x)$ and let $\mathsf{Q}_\exists$ be the number of $\exists$-quantifiers in Q, i.e., $\mathsf{Q}_\exists = H(\mathsf{Q})$ where

$$\begin{aligned} H([]) &= 0, \\ H([\ulcorner \exists \mathsf{v}_i \urcorner] ^\frown x) &= \mathsf{S}(H(x)) \quad \text{and} \\ H([\ulcorner \forall \mathsf{v}_i \urcorner] ^\frown x) &= H(x). \end{aligned}$$

Definition 5.12. A function $\Psi : \mathfrak{M} \mapsto \mathfrak{M}$ is a *Skolem operator* for the sequence of quantifiers Q if

$$\forall x, y \forall j < \mathsf{Q}_\exists \Big(\forall i < f_{\mathsf{Q}}(\mathsf{S}(j)) \big[(x)_i = (y)_i \big] \rightarrow (\Psi(x))_j = (\Psi(y))_j \Big).$$

If φ is a ${}^*\mathscr{L}_{\mathfrak{M}}$-formula and Q is a sequence of quantifiers then let $\mathsf{Q}\varphi$ denote the formula we get by preceding φ with the quantifiers Q.[4]

Define two $\mathtt{PA}$-definable functions $G^\forall, G^\exists : \mathfrak{M} \mapsto \mathfrak{M}$ such that the following is provable in $\mathtt{PA}$:

$$\begin{aligned} G^\exists([], y) = G^\forall([], y) &= 0 \\ G^\forall([\ulcorner \forall \mathsf{v}_i \urcorner] ^\frown x, 0) &= i \\ G^\exists([\ulcorner \exists \mathsf{v}_i \urcorner] ^\frown x, 0) &= i \\ G^\forall([\ulcorner \exists \mathsf{v}_i \urcorner] ^\frown x, y) &= G^\forall(x, y) \\ G^\exists([\ulcorner \exists \mathsf{v}_i \urcorner] ^\frown x, \mathsf{S}(y)) &= G^\exists(x, y) \\ G^\forall([\ulcorner \forall \mathsf{v}_i \urcorner] ^\frown x, \mathsf{S}(y)) &= G^\forall(x, y) \\ G^\exists([\ulcorner \forall \mathsf{v}_i \urcorner] ^\frown x, y) &= G^\exists(x, y). \end{aligned}$$

[4] The exact definition of $\mathsf{Q}\varphi$ depends on the Gödel numbering.

Also, define $g^{\forall}_{\mathsf{Q}}(x) = G^{\forall}(\mathsf{Q}, x)$ and $g^{\exists}_{\mathsf{Q}}(x) = G^{\exists}(\mathsf{Q}, x)$. Informally, if q is either $\forall$ or $\exists$ then $g^{q}_{\mathsf{Q}}(x) = i$ if the $(x+1)$st q-quantifier in Q bounds the variable v_i.

If $\mathsf{Q}\varphi$ is a sentence and Ψ is a Skolem operator for Q then $\varphi[\Psi, a]$ will denote the ${}^{*}\mathscr{L}_{\mathfrak{M}}$-sentence we get by substituting v_i, where

$$i = (\mathrm{M}x)\big[g^{\forall}_{\mathsf{Q}}(x) = l\big],$$

by $(a)_l$ and v_j, where

$$j = (\mathrm{M}x)\big[g^{\exists}_{\mathsf{Q}}(x) = l\big],$$

by $(\Psi(a))_l$, in φ. Here $(\mathrm{M}x)\varphi(x)$ means 'the greatest x such that $\varphi(x)$,' we are using it instead of (μx) to take care of situations like

$$\forall \mathsf{v}_0 \exists \mathsf{v}_1 \forall \mathsf{v}_0 (\mathsf{v}_0 = \mathsf{v}_1)$$

where it is the second $\forall$ quantifier bounding v_0, not the first.

If Ψ is a Skolem operator such that the function $\Psi : \mathfrak{M} \mapsto \mathfrak{M}$ is $\mathscr{L}_{\mathfrak{M}}$-definable then we say that Ψ is a *definable Skolem operator*.

Definition 5.13 ([Kra76]). A satisfaction class Σ is *complete with respect to definable Skolem operators* if

$$\mathsf{Q}\varphi \in \Sigma \quad \text{iff} \quad \text{there exists a definable Skolem operator } \Psi \text{ for } \mathsf{Q} \text{ such that } \varphi[\Psi, a] \in \Sigma \text{ for all } a \in \mathfrak{M}.$$

Remark 5.14. In the standard model $\mathbb{N}$ of PA the only satisfaction class,

$$\Sigma_0 = \{\, \ulcorner\varphi\urcorner \mid \varphi \in \mathrm{ElDiag}(\mathbb{N}) \,\},$$

is complete with respect to definable Skolem operators. This follows from the more general fact that in any $\mathfrak{M} \models \mathtt{PA}$ and $\mathscr{L}_{\mathfrak{M}}$-sentence $\mathsf{Q}\varphi$ we have

$$\mathfrak{M} \models \mathsf{Q}\varphi \quad \text{iff} \quad \text{there exists a definable Skolem operator for } \mathsf{Q} \text{ such that } \mathfrak{M} \models \varphi[\Psi, a] \text{ for all } a \in \mathfrak{M}.$$

Let us define the corresponding rule:

$$\frac{\ldots\ \Gamma, \varphi[\Psi, a]\ \ldots\ {}_{a \in \mathfrak{M}}}{\Gamma, \mathsf{Q}\varphi} \tag{Skolem-rule}$$

where Ψ is a definable Skolem operator for the sequence of quantifiers Q.

Proposition 5.15. *A maximally consistent set in $\mathfrak{M}$-logic with Skolem-rule is a satisfaction class complete with respect to definable Skolem operators.*

Proof. Left to the reader. $\square$

5.5 Closure under predicate logic

Let $\mathtt{Prf}_\sigma(y)$ denote the formula expressing that there is a (nonstandard) predicate logic proof of y from hypothesis satisfying $\sigma(x)$. The exact definition of the formula is left to the reader to figure out. We will write $\Lambda \vdash^* \varphi$ for

$$\mathfrak{M} \models \mathtt{Prf}_{x\in\Lambda}(\varphi).$$

We will add the rule

$$\frac{\dots\ \Lambda, \varphi_i\ \dots\ {}_{i<a}}{\Gamma} \qquad \text{if } \Big\{ \bigvee \Lambda \vee \varphi_0, \dots, \bigvee \Lambda \vee \varphi_{a-1} \Big\} \vdash^* \bigvee \Gamma. \tag{Pred}$$

to $\mathfrak{M}$-logic.

Proposition 5.16. *Satisfaction classes closed under nonstandard first-order provability are precisely the maximally consistent sets in $\mathfrak{M}$-logic with Pred.*

Proof. Left to the reader. □

Let $\mathtt{PA}^*$ be the set of all standard and nonstandard instances of the axioms of $\mathtt{PA}$.

Theorem 5.17 ([Kot85]). *If Σ is a satisfaction class such that*

$$\langle \mathfrak{M}, \Sigma \rangle \models \forall x \big(\mathtt{Sent}(x) \wedge \mathtt{Prf}_{x\in\mathtt{PA}^*}(x) \to x \in \Sigma \big)$$

then $\langle \mathfrak{M}, \Sigma \rangle$ satisfies Δ_0-induction, i.e.,

$$\langle \mathfrak{M}, \Sigma \rangle \models \varphi(0, \Sigma) \wedge \forall x \big(\varphi(x, \Sigma) \to \varphi(\mathsf{S}(x), \Sigma) \big) \to \forall x\, \varphi(x, \Sigma)$$

for every $\mathscr{L}_{\mathfrak{M}} \cup \{ \Sigma \}$-formula $\varphi(x, \Sigma)$ which is Δ_0.

Remark 5.18. The arithmetical part of any $\langle \mathfrak{M}, \Sigma \rangle$ satisfying the condition in the theorem is stronger than $\mathtt{PA}$; for example, the consistency of $\mathtt{PA}$ is provable in such a model, since otherwise $0 \neq 0 \in \Sigma$.

Remark 5.19. The previous remark shows that some countable recursively saturated models of $\mathtt{PA}$ does *not* admit satisfaction classes closed under nonstandard provability in $\mathtt{PA}$. It might still be the case that any countable recursively saturated model of $\mathtt{PA}$ admits a satisfaction class closed under nonstandard first-order provability since it may fail to include $\mathtt{PA}^*$.

6

Conclusion and further work

In the first and second chapter we gave some background information, including a short historical survey of the study of nonstandard truth. In Chapter 3 we introduced a new definition of satisfaction class in a language with function symbols. We also discussed the drawback of defining a satisfaction class as a set of pairs of formulas and elements. The main part of the chapter led up to Theorem 3.55, some applications were presented in the end of the chapter, such as the existence of satisfaction classes making the sentence $\mathsf{S}(\mathsf{S}(\ldots(\mathsf{S}(0))\ldots)) = \mathsf{c}_a$ true for any nonstandard number of S symbols and any nonstandard a. Our definitions of $\mathfrak{M}$-logic and template logic is rather different from other authors and we think our notions is easier to work with.

In the chapter that followed we introduced *free* satisfaction class, it is a weaker notion than satisfaction classes and in some sense it is a more natural notion, e.g., free $\mathfrak{M}$-logic is more natural than $\mathfrak{M}$-logic. We proved one characterisation of free satisfaction classes in terms of free $\mathfrak{M}$-logic.

Chapter 5 presented some ideas of how to remove pathologies. One famous pathology is $0 \neq 0 \vee \ldots \vee 0 \neq 0$ which can be made true for any nonstandard number of repetitions. We highlighted some other pathologies and gave ideas of how to remove those. We stated the question of whether there are satisfaction classes closed under nonstandard propositional (or predicate) proofs in any countable recursively saturated model.

We end this chapter be listing the open questions stated in the thesis.

Conjecture 3.1. *There is a satisfaction class Σ, in the sense of [Kra76] and [KKL81], such that $\langle \epsilon_a, [0] \rangle \in \Sigma$ and $\langle \epsilon'_a, [0] \rangle \notin \Sigma$ for some $a \in \mathfrak{M}$.*

Remember that ϵ_0 is $\mathsf{v}_0 \neq \mathsf{v}_0$ and ϵ_{i+1} is $\epsilon_i \vee \epsilon_i$; ϵ'_0 is $\mathsf{v}_0 \neq 0$ and ϵ'_{i+1} is $\epsilon'_i \vee \epsilon'_i$. We think that by redefining $\mathfrak{M}$-logic to work with pairs of $\mathscr{L}_A$-formulas and elements of $\mathfrak{M}$ it should be possible to prove the conjecture.

Question 3.7. *Are satisfaction classes built up by two "parts;" one with closed terms and equality and one with "the rest," in the following sense: Given a relation $\sim$ satisfying the conditions in Proposition 3.6 is there a satisfaction class Σ such* *that the relation $\sim_\Sigma$ coincides with $\sim$?*

A partial answer is given in Proposition 3.58. The general question seems to be hard, since it is a question of whether a set of equalities is consistent or not in $\mathfrak{M}$-logic. See also Question 5.2.

Question 4.6. *Which $\mathscr{L}_A$-structures $\mathfrak{N} \supseteq \mathfrak{M}$ are $\mathfrak{M}_\Sigma$ for some free satisfaction class Σ?*

It is a natural question to ask. It might be the case that the structures $\mathfrak{M}_\Sigma$ have very specific properties, analogue to the models arising in the arithmetised completeness theorem.

Question 4.7. *Are there free satisfaction classes Σ, which are not satisfaction classes, cuh that the canonical map $f : \mathfrak{M} \to \mathfrak{M}_\Sigma$ is an elementary embedding?*

If this is true, is there a corresponding extension of free $\mathfrak{M}$-logic?

Question 5.2. *How should Figure 5.1 be completed?*

The question marks are all in the column 'Consistent?', thus this is a question of proving consistency of extensions of $\mathfrak{M}$-logic. The only tool we have to do so is Theorem 3.55 but it will not help us in this situation. The only plausible approach we have found is to alter the definition of template logic and template structure, but every attempt of this has ended with tears. We think this question is very hard.

Bibliography

[Ben99] Ermanno Bencivenga. Free logics. In *Handbook of philosophical logic volume III*, pages 373–426. Kluwer, Dordrecht, 1999.

[BS76] Jon Barwise and John Schlipf. An introduction to recursively saturated and resplendent models. *J. Symbolic Logic*, 41(2):531–536, 1976.

[HP98] Petr Hájek and Pavel Pudlák. *Metamathematics of first-order arithmetic*. Springer-Verlag, Berlin, 1998. Second printing.

[Kay91] Richard Kaye. *Models of Peano arithmetic*. The Clarendon Press Oxford University Press, New York, 1991. Oxford Science Publications.

[KKL81] H. Kotlarski, S. Krajewski, and A. H. Lachlan. Construction of satisfaction classes for nonstandard models. *Canad. Math. Bull.*, 24(3):283–293, 1981.

[Kle52] S. C. Kleene. Finite axiomatizability of theories in the predicate calculus using additional predicate symbols. Two papers on the predicate calculus. *Mem. Amer. Math. Soc.*, 1952(10):27–68, 1952.

[Kot85] Henryk Kotlarski. Bounded induction and satisfaction classes. In *Proceedings of the third Easter conference on model theory (Gross Köris, 1985)*, pages 143–167. Humboldt Univ. Berlin, 1985.

[KR90a] Henryk Kotlarski and Zygmunt Ratajczyk. Inductive full satisfaction classes. *Ann. Pure Appl. Logic*, 47(3):199–223, 1990.

[KR90b] Henryk Kotlarski and Zygmunt Ratajczyk. More on induction in the language with a satisfaction class. *Z. Math. Logik Grundlag. Math.*, 36(5):441–454, 1990.

[Kra76] S. Krajewski. Non-standard satisfaction classes. In *Set theory and hierarchy theory (Proc. Second Conf., Bierutowice, 1975)*, pages 121–144. Lecture Notes in Math., Vol. 537. Springer, Berlin, 1976.

[Lac81] A. H. Lachlan. Full satisfaction classes and recursive saturation. *Canad. Math. Bull.*, 24(3):295–297, 1981.

[Lam01] Karel Lambert. Free logics. In *The Blackwell guide to philosophical logic*, pages 258–279. Blackwell, Oxford, 2001.

[Men97] Elliott Mendelson. *Introduction to mathematical logic*. Chapman & Hall, London, fourth edition, 1997.

[Rob63] Abraham Robinson. On languages which are based on non-standard arithmetic. *Nagoya Math. J.*, 22:83–117, 1963.

[Smi84] Stuart Thomas Smith. *Nonstandard syntax and semantics and full satisfaction classes for models of arithmetic*. PhD thesis, Yale University, 1984.

[Twe00] Ian Tweddle. *Simson on porisms*. Springer-Verlag London Ltd., London, 2000. An annotated translation of Robert Simson's posthumous treatise on porisms and other items on this subject.