View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by University of Birmingham Research Archive, E-theses Repository

Truth, Collection and Deflationism in Models of Peano Arithmetic

by

Alexander Marcus Jones

A thesis submitted to the University of Birmingham for the degree of Master of Research

School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
February 2015

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis focuses on adding collection axioms to satisfaction classes and exploring the suitability of a formal deflationary truth predicate. Chapter 2 proves that every nonstandard, recursively saturated model of PA has a satisfaction class in which all collection axioms are true. Chapter 3 explores collection axioms for the language with the satisfaction predicate, $\mathscr{L}_S$, and proves that these entail the theory of chapter 2. This chapter then demonstrates a method of closing a model with a satisfaction class to produce a new model with an induced satisfaction class, which it is conjectured will not satisfy all Σ_1 collection axioms in $\mathscr{L}_S$. In chapter 4 we conjecture that a new formulation of Visser and Enayat's construction of extensions of models with a satisfaction classes [5] will provide elementary extensions. Using this conjecture, we demonstrate new Tarski axioms provide satisfaction classes with Σ_1 collection axioms and that these axioms can be built into the theory by reducing the language to one where formulas are stratified. Finally, in chapter 5 we argue for a new definition of a deflationary truth predicate and show that this entails there are no formalisations of a deflationary truth predicate for the full nonstandard language of arithmetic.

Acknowledgements

First and foremost my sincere thanks and acknowledgements go to my supervisor Dr. Richard Kaye. His encouragement guided me towards this area and throughout this project his tireless assistance, advice and knowledge has been invaluable. I must also thank the staff and students at the University of Birmingham, and in particular the mathematics department, for creating such a friendly and supportive place to study. Studying my M.Res here has been a pleasure. Finally, to all of my friends and family who have kept me going whilst writing this thesis, a huge thank you for all the encouragement you provided, even without understanding the content of this thesis!

Contents

Chapter 1

Introduction

Truth is a topic that is of interest to both mathematics and philosophy. In philosophy research attempts to understand what truth is. The aim is to say what it means for a sentence to be true and to describe when a sentence is true. In mathematics, research attempts to formalise conceptions of truth. The mathematician aims to provide a formula that states when a sentence is true or false and to explore the consequences of this formula. These aims are interconnected and often inform one another; to formalise truth one needs to at least partially understand truth and an understanding of truth can be informed by a formalisation and its consequences. The aim of this thesis is to provide new formalisations of truth for arithmetic and explore the consequences of this, in particular its consequences for philosophical conceptions of truth.

1.1 Truth in Arithmetic

A natural place to attempt to formalise truth is arithmetic. Arithmetic is very well understood and has the very natural first order formalisation of Peano Arithmetic. This formal system consists of the the rules of predicate logic, together with basic axioms of arithmetic and the axiom schema of induction. Peano Arithmetic (PA) has been studied a great deal. It is therefore well comprehended by mathematicians and a variety of formal tools have been developed to do this. Further, its language, the language of arithmetic, is a language which can be formalised and expressed within PA itself. Therefore Peano Arithmetic is sufficiently powerful to talk about its own sentences. This means there is hope that truth can be formalised for this theory and that its consequences can be formally studied and understood.

Tarski [20] was the first mathematician to make real progress with formalising truth for arithmetic. He developed the formal syntax to define the truth of atomic sentences and then provided axioms to define truth for sentences built from connectives and quantifiers. He also proved one of the most important theorems in the field, that truth cannot be defined within the language it defines truth for. This has the two corollaries that there is no definition of truth for arithmetic in the language of arithmetic and a truth predicate cannot talk about the truth of all sentences involving the word 'true'.

Tarski's definition of truth can be applied to all finite sentences of arithmetic by inductively applying the axioms to the sentence. This induction is not done using the induction schema of PA, because this schema only applies to formulas of arithmetic and by Tarski's theorem the truth predicate cannot be defined in arithmetic. Whilst this is unproblematic for finite sentences, these are not the only sentences that can be discussed by Peano Arithmetic.

The intended model of PA, it can be argued, is the structure of the natural numbers with the usual constants 0 and 1 and the usual interpretation of $+$ and $\cdot$. Because of this, this is often called the standard model. Peano Arithmetic has many other models, both countable and uncountable, however, which contain so-called 'nonstandard' numbers, which are colloquially infinitely large. PA interprets some of these numbers as expressing sentences in the language of arithmetic, however these sentences have length which is interpreted by us as infinitely large. Tarski's axioms for truth are not able to definite truth for these sentences, because the induction cannot be applied infinitely many times.

For the theory of PA, a theory of truth that cannot define truth for all sentences of arithmetic, even if we would not recognise most of these sentences as real, is incomplete. Therefore, attempts have been made to provide new definitions of truth for these nonstandard sentences. Krajewski [14] made the natural adjustment to specify that the truth definition applies to all nonstandard sentences and ensures that all the Tarski axioms are satisfied. This theory becomes rich with mathematical structure and has the unattractive consequence that the truth of intuitively obvious nonstandard sentences may not conform to expectation.

There have been several attempts to improve this theory to restrict these unattractive consequences, the most common one being to add induction axioms to the language with the truth predicate as well. This creates a very strong mathematical theory that is interesting as it is able to prove the consistency of Peano Arithmetic, something that cannot be done within PA by Gödel's Second Incompleteness Theorem. This means that many models of PA cannot use this as a definition of truth.

1.2 Philosophical Links

Philosophical analysis has established understanding of the concept of truth. Truth is a property of sentences themselves, not of the content of the sentence. Further, truth is understood as a general concept that applies to many different types of sentence. For example, arithmetical truth has the same quality as scientific truth or literary truth, even if the methods of discovering these truths differ. One property of truth that is agreed is that for all propositions p, the sentence expressing p is true if and only if p. This is known as the equivalence schema and is an important criterion for a mathematical formalisation of truth.

One conception of truth, deflationism, goes beyond this statement and says that the equivalence schema is all that is understood by truth and it has no other properties. This is in contrast to a substantial theory of truth which offers further properties of truth with ontological commitments. For example, a theory of truth may state that if a sentence expressing a proposition p is true, then p corresponds to an actual state of affairs. A deflationary conception of truth proposes interesting consequences for a formalisation of truth and appears to specify that a theory of truth must be provided for models of arithmetic which model that PA is inconsistent. If truth could not be formalised for these models, then it suggests that truth has properties not recognised by the deflationist.

1.3 Areas of Inquiry

It is therefore of both mathematical and philosophical interest to establish formalisations of truth for nonstandard models of Peano Arithmetic and this is what shall be explored within this thesis. We will work from the standard theory of satisfaction classes, as posed by Krajewski and research alternative adjustments to the theory. These adjustments will have interesting consequences for the deflationary theory of truth and it is a secondary aim to explore what these consequences will be.

Our avenue of approach to adjusting the theory of satisfaction classes is to consider the axiom schema of collection, rather than induction. The collection axioms are entailed by the induction axioms, but are strictly weaker than them. They can be very weak in their consequences, but in certain circumstances are also very useful. At the very least the collection axioms added to satisfaction classes will create a theory of mathematical interest.

Within Chapter 2 we shall first demonstrate how a satisfaction class can be built for any nonstandard recursively saturated model which makes all the collection axioms true. We shall then, in Chapter 3, look at adding collection axioms in the language with the satisfaction class and explore closures of models and a satisfaction class to obtain models without full collection. In Chapter 4 we will then explore obtaining end extensions of models with a satisfaction class and the collection axioms that can be obtained there. Finally, in Chapter 5 we explore what this means for a deflationary theory of truth and argue that a formalised deflationary theory of truth must consider a restricted language of arithmetic.

Within this thesis we will use standard notation and theorems from models of arithmetic, as in the literature. For more details, the reader is referred to Appendix A. The reader unfamiliar with the collection axioms is referred to Appendix D for a thorough summary of the literature on these, as well as new results linking collection to tall and short recursively saturated models. The reader unfamiliar with satisfaction classes is referred to Appendix E for a thorough summary of these, as well as the original consideration of satisfaction classes with a relational interpretation of negation. Finally, a summary of deflationary theories of truth and one attempt in the literature to formalise them is given in Appendix C.

Chapter 2

Satisfying the Collection Axioms

In this chapter we focus on ensuring that all the collection axioms are satisfied, i.e. that $S \models \text{Coll}$, in the sense of Notation E.2.5. We prove that satisfaction classes of this type exist by working in a variant of M-Logic, which is defined in Section E.3, which has an added rule which behaves in the same manner as collection. To start with, we define this new form of logic, which we shall call M-C-Logic, which is specific to a given model $M \models \text{PA}$.

2.1 M-C-Logic

M-C-Logic, as a variant of M-Logic, is a formal system in the language ${}^*\mathscr{L}_A(M)$ and works with the same sequents of M-Logic. The difference between these two systems is in the extra strength of M-C-Logic, which we define below.

Definition 2.1.1 (Provability)**.** *We say that* $\vdash_{MC} \Delta$*, where* Δ *is a sequent, if there is a proof of the disjunction of all sentences of* Δ *from the rules of* M*-C-Logic. These are the structural and provability rules of* M*-Logic, with the additional C-Rule. Let* Δ *be a sequent,* $\theta(x, y)$ *be a* ${}^*\mathscr{L}_A(M)$*-formula and* t *a term. The C-Rule is:*

15. *(C-Rule) If* $\vdash_{MC} \Delta, \neg\forall x{<}t\exists y{<}s\theta(x, y)$ *for all terms* s*,*
    then $\vdash_{MC} \Delta, \neg\forall x{<}t\exists y\theta(x, y)$

We see that the C-Rule in M-C-Logic behaves in the same way as we expect collection axioms to. We prove this formally, using the provability rules of M-C-Logic, below.

Lemma 2.1.2. $\forall x{<}t\exists y\theta(x, y) \vdash_{MC} \exists z\forall x{<}t\exists y{<}z\theta(x, y)$.

Proof. We have by Rule 4 and weakening that:

$$\vdash_{MC} \neg\forall x{<}t\exists y\theta(x, y), \forall x{<}t\exists y{<}s\theta(x, y), \neg\forall x{<}t\exists y{<}s\theta(x, y)$$

for all terms s. Thus, by M-Rule a and the C-Rule, we have that:

$$\vdash_{MC} \neg\forall x{<}t\exists y\theta(x, y), \exists z\forall x{<}t\exists y{<}z\theta(x, y), \neg\forall x{<}t\exists y\theta(x, y).$$

Hence, using structural rules 2 and 3 we have that:

$$\vdash_{MC} \neg\forall x{<}t\exists y\theta(x, y), \exists z\forall x{<}t\exists y{<}z\theta(x, y)$$

as desired. □

We shall give M-C-Logic the following semantic interpretation.

Definition 2.1.3. *We say that a full satisfaction class S over M is a full collection satisfaction class if $(M,S) \vDash S(\ulcorner \neg\forall x{<}t\exists y{<}s\theta(x,y)\urcorner, a)$ for all s if and only if $(M,S) \vDash S(\ulcorner \neg\forall x{<}t\exists y\theta(x,y)\urcorner, a)$. We write $\vDash_{MC} \Delta$ to mean that every full collection satisfaction class S over M,makes some $\delta \in \Delta$ true.*

Therefore if M-C-Logic is sound, complete and consistent for a given model M, then we know that M has a satisfaction class S which believes all collection axioms are satisfied. We prove this below.

Lemma 2.1.4. *If $M \vDash \mathrm{PA}$ such that M-C-Logic is sound, complete and consistent, then $(M,S) \vDash S(\ulcorner \mathrm{Coll}_\theta \urcorner, a)$ for all (potentially nonstandard) formulae $\theta(x,y)$.*

Proof. Suppose $M \vDash \mathrm{PA}$ such that M-C-Logic is sound, complete and consistent and $\theta(x,y)$ be a formula. From the semantic interpretation of M-C-Logic, we know that $\vDash_{MC} \alpha(t), \beta(t)$ means that $(M,S) \vDash S(\ulcorner \alpha(t) \urcorner, a)$ or $S(\ulcorner \beta(t) \urcorner, a)$. We also know that Coll_θ is of the form $\alpha(t) \to \beta(t)$ where $\alpha(t)$ is $\forall x{<}t\exists y\theta(x,y)$ and $\beta(t)$ is $\exists z\forall x{<}t\exists y{<}z\theta(x,y)$. We know by definition that $S(\ulcorner \alpha \to \beta \urcorner, a)$ if and only if $\neg S(\ulcorner \alpha \urcorner, a) \vee S(\ulcorner \beta \urcorner, a)$. By Lemma 2.1.2 we know that $\vdash_{MC} \neg\alpha, \beta$. Thus, we have that $(M,S) \vDash \neg S(\ulcorner \alpha \urcorner, a) \vee S(\ulcorner \beta \urcorner, a)$ and we are done. □

We can thus use M-C-Logic to build satisfaction classes which establish truth for the collection axioms. We do this explicitly now in a similar manner to the last chapter, by proving that, for models which are countable and recursively saturated, M-C-Logic is both sound, complete and consistent. We start off by showing that M-C-Logic is sound.

Theorem 2.1.5 (Soundness). *If* $\vdash_{MC} \Delta$ *for any sequent* Δ*, then* $\vDash_{MC} \Delta$*.*

Proof. We saw in the proof of Theorem E.3.10 that M-Logic is sound. Thus, we modify the proof of that theorem, by showing that the addition of the C-Rule preserves soundness:

If the last derivation used the C-Rule to get $\vdash_{MC} \Delta, \neg\forall x{<}t\exists y\theta(x,y)$ from $\vdash_{MC} \Delta, \neg\forall x{<}t\exists y{<}s\theta(x,y)$, then by our inductive hypothesis every full collection satisfaction class over M makes some $\delta \in \Delta$ true or $\neg\forall x{<}t\exists y{<}s\theta(x,y)$ true. In the latter case, we have by our additional condition for a collection satisfaction class that the satisfaction class makes $\neg\forall x{<}t\exists y\theta(x,y)$ true. Hence every full collection satisfaction class over M makes some $\delta \in \Delta \cup \{\neg\forall x{<}t\exists y\theta(x,y)\}$ true. □

We now show that, for a countable model M, M-C-Logic is complete, by constructing a suitable satisfaction class.

Theorem 2.1.6 (Completeness). *Let* M *be countable. If* $\vDash_{MC} \Delta$ *for any sequent* Δ*, then* $\vdash_{MC} \Delta$*.*

Proof. We follow the proof of Theorem E.3.12 here and show that $\Gamma \nvdash_{MC} \varnothing$ implies that there is a full collection satisfaction class S over M making each $\gamma \in \Gamma$ false. To do this we again inductively construct an infinite set $\Sigma = \bigcup_{i\in\mathbb{N}} \Sigma_i$ where $\Sigma_0 = \Gamma$ and each Σ_i is defined at some stage by dovetailing the following processes:

- Given Σ_i take $a \in M$ and a formula $\phi(x)$. If $\Sigma_i, \phi(a) \nvdash_{MC} \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\phi(a)\}$, else take $\Sigma_{i+1} = \Sigma_i$.

- Given Σ_i take $a \in M$ and a formula $\phi(x)$. If $\Sigma_i, \neg\phi(a) \nvdash_{MC} \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\neg\phi(a)\}$, else take $\Sigma_{i+1} = \Sigma_i$.

- Given Σ_i take a sentence σ. If $\Sigma_i, \sigma \nvdash_{MC} \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\sigma\}$, else take $\Sigma_{i+1} = \Sigma_i \cup \{\neg\sigma\}$.

- Given Σ_i take $a, t \in M$ and a formula $\theta(x, y)$. If we have that: $\Sigma_i, \forall x{<}t \exists y{<}a \theta(x, y) \nvdash_{MC} \varnothing$, then $\Sigma_{i+1} = \Sigma_i \cup \{\forall x{<}t \exists y{<}a \theta(x, y)\}$. Else, we take $\Sigma_{i+1} = \Sigma_i \cup \{\neg\forall x{<}t \exists y \theta(x, y)\}$.

We can perform this last step, since given any finite set $\Gamma \nvdash_{MC} \varnothing$ we have that $\Gamma, \forall x{<}t \exists y{<}a \theta(x, y) \nvdash_{MC} \varnothing$ for some a or that $\Gamma \vdash_{MC} \neg\forall x{<}t \exists y \theta(x, y)$. This is from $\Gamma, \forall x{<}t \exists y{<}a \theta(x, y) \vdash_{MC} \varnothing$ implies $\Gamma \vdash_{MC} \neg\forall x{<}t \exists y{<}a \theta(x, y)$, by the syntax of M-C-Logic and thus by the C-Rule $\Gamma \vdash_{MC} \neg\forall x{<}t \exists y \theta(x, y)$.

We now again enumerate all ${}^\star\mathscr{L}_A$-sentences, formulas and elements of M to dovetail the above processes to construct Σ. We then know that Σ forms a satisfaction class by completing the steps of the proof of Theorem E.3.12. Finally, we know that Σ is a full collection satisfaction class by the C-Rule.

Thus, if $\Gamma \nvdash_{MC} \varnothing$, then $\Gamma \nvDash_{MC} \varnothing$. □

We thus require only that M-C-Logic is consistent, for us to have a satisfaction class that makes all collection axioms true. We prove this in the following subsection.

2.1.1 Consistency of M-C-Logic

To prove the consistency of M-C-Logic, we first require the following lemma, that models of collection and short recursive saturation satisfy collection over an infinite disjunction of statements, Coll$\bigvee$ as defined in Definition D.2.9.

Lemma 2.1.7. *If a model $M \models$ Coll and M is short recursively saturated, then $M \models$ Coll$\bigvee$.*

Proof. We prove this by contrapositive. Suppose M is a short recursively saturated model and that $M \models \neg\exists z \forall x{<}t \exists y{<}z \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$. Thus, we can rewrite this as:

$$M \models \forall z \exists x{<}t \forall y{<}z \bigwedge_{n\in\mathbb{N}} \neg\theta_n(x, y, \bar{a}).$$

This implies that $M \models \forall z \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \forall y{<}z \neg\theta_i(x, y, \bar{a})$. We can swap the universal quantification and conjunction at the beginning of this sentence, to get that $M \models \bigwedge_{n\in\mathbb{N}} \forall z \exists x{<}t \bigwedge_{i<n} \forall y{<}z \neg\theta_i(x, y, \bar{a})$. Now, by the contrapositive of Coll, we have that:

$$M \models \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \forall y \neg\theta_i(x, y, \bar{a}).$$

Now, using the fact that M is short recursively saturated we get that $M \models \exists x{<}t \bigwedge_{n\in\mathbb{N}} \forall y \neg\theta_n(x, y, \bar{a})$. Thus, by swapping the universal quantifier and the conjunction and by negation, we get that $M \models \neg\forall x{<}t \exists y \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$. Therefore, by contrapositive, $M \models$ Coll$\bigvee$. □

We now prove the consistency of M-C-Logic, for a countable and recursively saturated model M. First we show that M-C-Logic is entailed by M-Logic with every instance of the collection axiom. Then we prove that M-Logic with every instance of the collection axiom is consistent, by showing that Diag(M) proves every instance of the collection axiom in FA-Logic.

We start by proving that M-C-Logic is no more powerful than M-Logic with the collection axioms.

Lemma 2.1.8. *If* $\Sigma \vdash_{MC} \Delta$ *for any sequents* Σ, Δ*, then for some collection axioms* $\text{Coll}_{\theta_1}, \text{Coll}_{\theta_2}, \ldots$ *we have* $\Sigma \cup \{\text{Coll}_{\theta_i} : i = 1, 2, \ldots\} \vdash_M \Delta$.

Proof. We show that, for any formula θ, if $\text{Coll}_\theta \vdash_M \neg\forall x{<}t\exists y{<}s\theta(x,y)$, then $\text{Coll}_\theta \vdash_M \neg\forall x{<}t\exists y\theta(x,y)$. In other words, we show that Coll_θ simulates the C-Rule for the formula θ. We rewrite Coll_θ as $\neg\alpha \vee \beta$ where α is $\neg\forall x{<}t\exists y\theta(x,y)$ and β is $\exists z\forall x{<}t\exists y{<}z\theta(x,y)$. We first use M-Rule a to get that $\text{Coll}_\theta \vdash_M \neg\exists z\forall x{<}t\exists y{<}z\theta(x,y)$. We rewrite this using the syntax of M-Logic to get that $\vdash_M \neg(\neg\alpha \vee \beta), \neg\beta$. We then use the weakening rule on this to get that $\vdash_M \neg(\neg\alpha \vee \beta), \neg\beta, \neg\alpha$. We now use rule 4 twice to get that $\vdash_M \beta, \neg\beta, \alpha, \neg\alpha$. Thus, by rules 8 and 9 we have that $\vdash_M \neg(\neg\alpha \vee \beta), \beta, \neg\alpha$ which allows us to use the cut rule to get that $\vdash_M \neg(\neg\alpha \vee \beta), \neg\alpha$, i.e. that $\text{Coll}_\theta \vdash_M \neg\forall x{<}t\exists y\theta(x,y)$. □

This tells us that the C-Rule in M-C-Logic is replicated by the first order collection schema and allows us an equivalent method of proof to show that M-C-Logic is consistent. If $\vdash_{MC} \varnothing$, then by Theorem 2.1.8 we know that $\text{Coll} \vdash_M \varnothing$. We will prove that this does not hold, and thus by contrapositive M-C-Logic is consistent.

Lemma 2.1.9. *Let* $M \vDash \text{PA}$ *be countable, recursively saturated and* Coll_θ *be a collection axiom. If* $\text{Diag}(M), \text{Coll}_\theta \vdash_{\text{FA}} \Delta$ *for any sequent* Δ*, then* $\text{Diag}(M) \vdash_{\text{FA}} \Delta$.

Proof Sketch. To prove this, we use the method of enumerating proofs of FA-Logic as in the proof of Theorem E.4.5. We then show that $\text{Diag}(M)$

can mimic the collection axioms, i.e. if $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \Delta, \forall x{<}t \exists y \theta(x, y)$, then $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \Delta, \exists z \forall x{<}t \exists y{<}z \theta(x, y)$. Therefore, to prove this, we suppose that $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \Delta, \forall x{<}t \exists y \theta(x, y)$. We rewrite this as:

$$\bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \forall x{<}t \exists y \theta(x, y).$$

Suppose that $\neg \forall a{<}t \bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \exists y \theta(a, y)$. Then we have that $\exists a{<}t \bigwedge_p \mathrm{Diag}(M) \nvdash^p_{\mathrm{FA}} \Delta, \exists y \theta(x, y)$. In other words there is some $a \in M$ such that $\mathrm{Diag}(M) \nvdash_{\mathrm{FA}} \Delta, \exists y \theta(a, y)$. However, it is an easy exercise to see that this contradicts with $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \forall x{<}t \exists y \theta(x, y)$ under the rules of FA-Logic. Therefore, we have that:

$$\forall a{<}t \bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \exists y \theta(a, y).$$

Now suppose that $\neg \forall a{<}t \exists c \bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \theta(a, c)$. Then we have that $\exists a{<}t \forall c \bigwedge_p \mathrm{Diag}(M) \nvdash^p_{\mathrm{FA}} \Delta, \theta(a, c)$. In other words for some $a \in M$ and for all $c \in M$ we have that $\mathrm{Diag}(M) \nvdash_{\mathrm{FA}} \theta(a, c)$. This is again a contradiction under our hypothesis and the rules of FA-Logic, so therefore we have that

$$\forall a{<}t \exists c \bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \theta(a, c).$$

By Lemma 2.1.7 if M is short recursively saturated, then it satisfies $\mathrm{Coll} \bigvee$. All recursively saturated models are short recursively saturated, therefore, we can use Coll $\bigvee$ to get that

$$\exists z \forall a{<}t \exists c{<}z \bigvee_p \mathrm{Diag}(M) \vdash^p_{\mathrm{FA}} \Delta, \theta(a,c).$$

Thus, we have that there is some proof in FA-Logic p which says given $z \in M$, then for all $a < t \in M$ there is $c < z \in M$ which gives us that $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \Delta, \theta(a,c)$. Therefore, by FA-Rules 13 and 14 we have that:

$$\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \Delta, \exists z \forall x{<}t \exists y{<}z \theta(x,y).$$

□

With this lemma in place, we are finally able to prove the consistency of M-C-Logic.

Theorem 2.1.10. *For $M \models \mathrm{PA}$ where M is countable, nonstandard and recursively saturated we have that M-C-Logic is consistent.*

Proof. If not, then we have that $\vdash_{MC} \varnothing$. However, this would mean that $\mathrm{Coll} \vdash_M \varnothing$ by Lemma 2.1.8. This in turn would entail that $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} \varnothing$ by Lemma 2.1.9, which is a contradiction. □

Thus, since M-C-Logic is complete, we know that we are able to build satisfaction classes which make every collection axiom true. This appears to have an interesting property, which differ to standard collection, which we look at now.

2.2 Satisfied Collection and Regular Collection

We know from Theorem D.1.10 that $\mathrm{PA}^{-} + \mathrm{I}\Delta_0 + \mathrm{Coll} \vdash \mathrm{PA}$. We also saw in Theorem E.5.16 that $S \models \mathrm{IND}$ if and only if $S \models \mathrm{LNP}$. In the latter theorem, we see that making induction axioms true in a satisfaction class appears to change nothing of their behaviour. The intuition becomes that if a full collection satisfaction class also has that $S \models \mathrm{I}\Delta_0$, then $S \models \mathrm{IND}$. Interestingly, this appears incorrect because the standard proof of Theorem D.1.10 does not run through for satisfaction classes.

The reason for this is because it requires a meta-induction on Σ_n formulae, where n ranges over nonstandard numbers also. We cannot do this within a normal satisfaction class, since we have no full axioms of induction for them, and an induction outside the model will not include these nonstandard numbers. This leads to the following conjecture.

Conjecture 2.2.1. *There are models $M \models \mathrm{PA}$ with a full satisfaction class S such that $S \models \mathrm{I}\Delta_0 + \mathrm{Coll}$ and $S \not\models \mathrm{IND}$.*

This is highly interesting and suggests that there might be a disconnect between ensuring sentences are true in a satisfaction class and having sentences as true from the perspective of the model. This prompts further research in this area, to see whether this disconnect is real and, if so, how widespread it is.

We can thus see that whilst we can ensure that all collection axioms are satisfied in a satisfaction class, this may not give us many of the interesting and useful properties we see in Chapter D that these axioms usually entail. Therefore, we now look for the stronger form of collection by adding the collection axioms to our expanded language $\mathscr{L}_S$.

Chapter 3

Collection in $\mathscr{L}_S$

In this chapter we explore adding the collection axioms to the language $\mathscr{L}_S$. We shall explore the consequences of adding these axioms to our theory, and then various methods of obtaining them.

3.1 $\mathrm{Coll}(S)$

We begin this endeavour by defining the theory of satisfaction classes with all collection axioms that we are now considering.

Definition 3.1.1 ($\mathrm{Coll}(S)$). *We denote by* $\mathrm{Coll}(S)$ *the theory* $\mathrm{PA} + S$ *is a full satisfaction class and* Coll_θ *holds for all formulas* θ *in the language with a satisfaction class* $\mathscr{L}_S$.

This theory differs to the previous theory considered, where each collection axiom is satisfied, since it ensures that we can use the collection axioms on formulas which contain instances of the satisfaction predicate. This ensures that if a model with a satisfaction class entails Coll(S), then the satisfaction class is a collection satisfaction class.

Theorem 3.1.2. *If* $(M, S) \models \text{PA} + S$ *is a full satisfaction class* $+\text{Coll}(S)$*, then* S *is a full collection satisfaction class.*

Proof. Suppose that $(M, S) \models S(\ulcorner \neg \forall x{<}t \exists y{<}s \theta(x, y) \urcorner, a)$ for all s. Then, we know by the Tarski Conditions that this holds if and only if $(M, S) \models \neg \forall a{<}f \exists b{<}g S(\ulcorner \theta(x, y) \urcorner, m)$ for all g and the suitable assignment m. This holds if and only if $(M, S) \models \neg \forall a{<}f \exists b S(\ulcorner \theta(x, y) \urcorner, m)$ by Coll(S). Therefore, this is if and only if $(M, S) \models S(\ulcorner \neg \forall x{<}t \exists y \theta(x, y) \urcorner, a)$ by the Tarski Conditions. □

The next question to ask, given Coll(S) entails the previous theory, is whether this is consistent, i.e. whether there are any models $(M, S) \models$ Coll(S). This is the case, as is shown in Appendices D and E.

Theorem 3.1.3. *If* $(M, S) \models \text{PA}(S)$*, then* $(M, S) \models \text{Coll}(S)$*.*

Proof. Let $(M, S) \models \text{PA}(S)$. By Theorem D.1.4 we know that if a model $M \models \text{PA}$ for some finite language $\mathscr{L} \supseteq \mathscr{L}_A$, then $M \models$ Coll. Therefore, since $\mathscr{L}_S \supseteq \mathscr{L}_A$ and $\mathscr{L}_S$ is finite, we have that $(M, S) \models \text{Coll}(S)$. □

It should be noted that there are also many models with a satisfaction class that do not satisfy Coll(S). Again, by using results from the previous chapters, it is easy to see that this is the case. We know that there are models of $\Delta_0 - \text{PA}(S)$ which are not models of PA(S). Therefore, these models do not satisfy Coll(S), else they would be models of PA(S) by Theorem D.1.10.

An interesting open question is whether there is an easier method of constructing models with a satisfaction class that do not satisfy Coll(S) and also do not satisfy $\Delta_0 - PA(S)$.

Question 3.1.4. *Is there a model* $M \models \mathrm{PA}$ *with a satisfaction class* S *such that* $(M, S) \not\models \mathrm{Coll}(S)$ *and* $(M, S) \not\models \Delta_0 - \mathrm{PA}(S)$.

Another interesting open question is whether we can construct models of Coll(S) which do not satisfy $\Delta_0 - \mathrm{PA}(S)$. This would be a desirable theory, as it would hopefully, unlike $\Delta_0 - \mathrm{PA}(S)$, not prove the consistency of PA.

Question 3.1.5. *Is there a model* $M \models \mathrm{PA}$ *with a satisfaction class* S *such that* $(M, S) \models \mathrm{Coll}(S)$ *and* $(M, S) \not\models \Delta_0 - \mathrm{PA}(S)$.

3.2 Closure of (M, S)

One way to approach Question 3.1.4 is to consider closures of a model with a satisfaction class. Classically, given $M \models \mathrm{PA} + \neg con(\mathrm{PA})$, one can take the model $\Sigma_1\text{-}cl_M(\varnothing)$, as defined in Definition A.2.4. It is then the case that $\Sigma_1\text{-}cl_M(\varnothing) \models \mathrm{I}\Delta_0 + \neg\mathrm{Coll}_{\Sigma_1}$ [12, Chapter 10]. This is a method to obtain models without full collection axioms and it is hoped that this method will generalise to satisfaction classes.

We shall provide a construction of a closure of a model with a satisfaction class that allows an induced satisfaction class to be defined on the closure. To do this we consider a model $(M, S) \models \mathrm{PA} + S$ is a full satisfaction class over M and work in the language $\mathscr{L}_A$ with the definable Skolem function f_{Skolem}.

Definition 3.2.1. *Let $A \subseteq M$. Then, $cl_{Skolem}(A)$ = the smallest set $B \supseteq A$ which is closed under 0, the successor function, the pairing function and the function f_{Skolem}. Here, we define*

$$f_{Skolem}(\theta(x, \bar{y}), a) = \begin{cases} \text{the least } x \text{ such that } S \models \theta(x, \bar{b}), & \\ \text{where } \bar{b} \text{ is assigned to the variables} & \text{if such an } x \text{ exists} \\ \bar{y} \text{ by } a & \\ 0 & \text{if no such } x \text{ exists} \end{cases}$$

This definition specifies that $cl_{\text{Skolem}}(A)$ is closed under 0 and the successor function to ensure that all natural numbers are contained within this domain. The reason that it is also closed under the pairing function is so that new numbers can be built, especially assignments, as required.

These criteria ensure that the closure, given $S \models$ LNP, is well behaved in a classical sense. It is an easy exercise to see that $cl_{\text{Skolem}}(A) = cl_M(cl_{\text{Skolem}}(A))$ in the usual definition of closure. Therefore, by Theorem A.2.5 we have that $A \subseteq cl_{\text{Skolem}}(A) \prec M$.

This closure is useful, in that it allows us to define a satisfaction class over it. We define this below, and prove that it is indeed a satisfaction class for $cl_{\text{Skolem}}(A)$, where $A \subset M$.

Definition 3.2.2. *Let $(M, S) \models$ PA + S is a full satisfaction class over M and work in the language $\mathscr{L}_A$ with the definable Skolem function f_{Skolem}. Suppose that $S \models$ IND and let N denote $cl_{Skolem}(A)$, where $A \subset M$. Define $S_{cl} \subseteq N \times N$ by $(N, S_N) \models S_N(\ulcorner\theta\urcorner, a)$ if and only if $(M, S) \models S(\ulcorner\theta\urcorner, a)$ and we have that $\ulcorner\theta\urcorner, a \in N$.*

Theorem 3.2.3. *Let $(M, S) \models \mathrm{PA} + S$ is a full satisfaction class over M and work in the language $\mathscr{L}_A$ with the definable Skolem function f_{Skolem}. Suppose that $S \models \mathrm{IND}$, $A \subset M$ and $N = cl_{Skolem}(A)$. Then S_{cl} as defined above is a satisfaction class over N.*

Proof. To prove this we check that each Tarski condition holds for any formula with Gödel-number in N. First, we note that since $S \models \mathrm{IND}$ we have by Theorem E.5.16 that $S \models \mathrm{LNP}$.

Clearly all atomic formulas $\theta(x)$, such that $\ulcorner\theta(x)\urcorner \in N$ are uniquely satisfied in N, so are true in S_{cl} for a suitable assignment $a \in N$.

If $(N, S_{cl}) \models S_{cl}(\ulcorner\alpha(x)\urcorner, a) \wedge S_{cl}(\ulcorner\beta(x)\urcorner, a)$, then we know that $(M, S) \models S(\ulcorner\alpha(x) \wedge \beta(x)\urcorner, a)$. Since $N \prec M$, and in particular N is a substructure, we have that $\ulcorner\alpha \wedge \beta\urcorner \in N$. Therefore $(N, S_{cl}) \models S_{cl}(\ulcorner\alpha \wedge \beta\urcorner, a)$.

If $(N, S_{cl}) \models S_{cl}(\ulcorner\alpha \wedge \beta\urcorner, a)$, then we know that $(M, S) \models S(\ulcorner\alpha(x)\urcorner, a) \wedge S(\ulcorner\beta(x)\urcorner, a)$. Since $N \prec M$, and in particular N is a substructure, we have that $\ulcorner\alpha\urcorner \in N$ and $\ulcorner\beta\urcorner \in N$. Therefore, we conclude that $(N, S_{cl}) \models S_{cl}(\ulcorner\alpha(x)\urcorner, a) \wedge S_{cl}(\ulcorner\beta(x)\urcorner, a)$

The case for Tarski Conditions 4) and 5) are proven similarly to the above.

Suppose $(N, S_{cl}) \models \exists b S_{cl}(\ulcorner\theta(x)\urcorner, a[^x/_b])$. Then, we have that $(M, S) \models S(\ulcorner\exists x\theta(x)\urcorner, a)$. Since $N \prec M$, and in particular N is a substructure, we have that $\ulcorner\exists x\theta(x)\urcorner \in N$ therefore $(N, S_{cl}) \models S_{cl}(\ulcorner\exists x\theta(x)\urcorner, a)$.

Now suppose $(N, S_{cl}) \models S_{cl}(\ulcorner\exists x\theta(x)\urcorner, a)$. Then, we have that $(M, S) \models \exists b S(\ulcorner\theta(x)\urcorner, a[^x/_b])$. We know that, since $S \models \mathrm{LNP}$, there is a least b such that $S \models \theta(b)$. Therefore, $b \in N$ and, since $\ulcorner\theta(x)\urcorner \in N$ we have that $(N, S_{cl}) \models \exists b S_{cl}(\ulcorner\theta(x)\urcorner, a[^x/_b])$.

The case for Tarski Condition 6) is proven similarly to the above. □

Therefore, we arrive at a closure of a model with a satisfaction class. This gives rise to the following conjecture that the above closure method behaves similarly to the classical result that $\Sigma_1\text{-}cl_M(\varnothing) \vDash \mathrm{I}\Delta_0 + \neg\mathrm{Coll}_{\Sigma_1}$.

Conjecture 3.2.4. *Let* $(M, S) \vDash \mathrm{PA} + S$ *is a full satisfaction class over* M*, where* $S \vDash \mathrm{IND}$*. Let* N *denote* $cl_{Skolem}(\mathbb{N})$ *and define* $S_{cl} \subseteq N \times N$ *by* $(N, S_{cl}) \vDash S_{cl}(\ulcorner\theta\urcorner, a)$ *if and only if* $(M, S) \vDash S(\ulcorner\theta\urcorner, a)$*. Then,* $(N, S_{cl}) \nvDash \mathrm{Coll}_{\Sigma_1}(S)$*.*

With this conjecture on how to obtain models without collection in $\mathscr{L}_S$, we now look at the far more interesting question as to how to obtain models with the collection axioms for $\mathscr{L}_S$. We begin this approach, by considering the techniques within the literature of Appendix D.

3.3 Results Which Follow From the Literature

In Appendix D we prove many interesting properties of the collection axioms from the literature. These theorems hold in a finite language $\mathscr{L}$, which is an extension of $\mathscr{L}_A$, as this means all of the results are applicable in $\mathscr{L}_S$ as well. We now list, without proof, some of the more relevant results that are entailed by this work.

Our first theorem comes from Theorem D.1.5 and proves a very general result that ω_1-like models satisfy collection.

Theorem 3.3.1. *If* $(M, S) \vDash \mathrm{PA} + S$ *is a full satisfaction class and* M *is* ω_1*-like, then* $(M, S) \vDash \mathrm{Coll}(S)$*.*

This result is not as general as one might hope, however, as Smith proved Theorem E.2.11 that not every uncountable model has a satisfaction class and no rather classless models have a satisfaction class. However, this result does tell us that there exist countable models of $\mathrm{Coll}(S)$ by the downward Löwenheim-Skolem theorem.

We can, in fact, do better than first order collection, by using Theorem D.3.2 to obtain models in which the second order collection axiom holds true in $\mathscr{L}_S$.

Theorem 3.3.2. *If κ is a regular cardinal, M is κ-like and $(M, S) \vDash \mathrm{PA} + S$ is a full satisfaction class, then $(M, S) \vDash \mathrm{Coll}^2(S)$.*

This may at first look to be a very powerful theorem, but we proved in Theorem D.3.6 that there is in fact only one countable model that does this, the standard model, and we have seen that this model has no satisfaction class. Further, we still have Smith's theorem that no rather classless uncountable models have satisfaction classes. Therefore, this theorem is not as strong as it could be.

One last result we can obtain is that at least one model of $\mathrm{Coll}^2(S)$ exists. We arrive at this via the proof of Theorem D.3.8.

Theorem 3.3.3. *There is a model of $\mathrm{PA}(S)$ which satisfies $\mathrm{Coll}^2(S)$.*

Thus, whilst these results provide us with full collection axioms in the language $\mathscr{L}_S$ for some uncountable models, they do not provide us with the theory $\mathrm{Coll}(S)$ for countable models. We now further explore the link between collection and end-extensions, to try and prove some results which also hold for countable models.

Chapter 4

Extensions of Models with Satisfaction Classes

Appendix D details that models of collection can be obtained by generating end-extensions of models. Thus, it seems natural that models of collection axioms in $\mathscr{L}_S$ could be obtained by generating an end-extension of models with a satisfaction class. We explore this thought in this chapter.

4.1 Visser-Enayat Construction

In their paper New Constructions of Satisfaction Classes [5] Visser and Enayat provide a method of generating an extension of a model with a partial satisfaction class. We shall denote this theorem using the notation of many-sorted logic.

Notation 4.1.1. *We write* $(M, F, S) \models \mathrm{PA} + S$ *is an* F *satisfaction class to mean the two-sorted logical structure which has domains* $M = (M, +, \cdot, <$

$, 0, 1, f_\theta, \ldots)$ *and* F*, relation* $S \subseteq F \times M$ *and function* $\ulcorner \cdot \urcorner : F \to M$*. We have that* $M \models \mathrm{PA}$ *and has all Skolem functions* f_θ *for each* $\theta \in F$*. Here* F *is a set of* ${}^\star\mathscr{L}_A$ *formulas, all of which can be identified in* M *using the function* $\ulcorner \cdot \urcorner$*, which is a suitable injective Gödel-numbering. Finally, we have that* S *is a partial satisfaction class over* M *for all formulas in* F*, using assignments coded in* M*.*

We write $(M, F_M, S_M) \subseteq (N, F_N, S_N)$ *to mean that* $M \prec N, F_M \subseteq F_N$ *and* $S_M = S_N$ *where the relation* S_N *is restricted to* $F_M \times M$*. We write* $(M, F_M, S_M) \subseteq_e (N, F_N, S_N)$ *to mean* $(M, F_M, S_M) \subseteq (N, F_N, S_N)$ *and* $M \prec_e N$*.*

We now state a slight generalisation of Visser and Enayat's theorem on construction of extensions of models with a satisfaction class using this many-sorted notation and provide a new sketch proof of it within this setup.

Theorem 4.1.2. *Suppose we have* $(M, F_M, S_M) \models \mathrm{PA} + S_M$ *is an* F_M *satisfaction class such that* F_M *is closed under immediate predecessor. Given a set of* ${}^\star\mathscr{L}_A$*-formulas* $F_N \supseteq F_M$ *such that we can injectively extend* $\ulcorner \cdot \urcorner$ *so that* $\ulcorner F_N \urcorner \subseteq M$ *and* F_N *is closed under immediate predecessor, there exists* N *and* S_N *such that* $(N, F_N, S_N) \models \mathrm{PA} + S_N$ *is an* F_N *satisfaction class and* $(M, F_M, S_M) \subseteq (N, F_N, S_N)$ *[5, Lemma 3.1].*

Proof Sketch. We build a theory $T \supseteq Th(M, F_M, S_M)$ and prove that it is consistent via compactness. We build T from each constant $a \in M$, and add each formula $\phi \in F_M$ to T and each formula $\psi \in F_N$ to T. We then add the sentence $\theta(\bar{a})$ for each $\mathscr{L}_A^{Skolem}$ sentence true of M, where $\bar{a} \in M$, in particular this ensures that we add all atomic sentences which hold in M. We then ensure $S(\theta, a)$ for all $\theta \in F_M$ and $a \in M$ such that $(M, F_M, S_M) \models S_M(\theta, a)$. We then, conversely, add $\neg S(\theta, a)$ for all $\theta \in F_M$ and $a \in M$ such that $(M, F_M, S_M) \models \neg S_M(\theta, a)$. We then establish that T makes all the Tarski Conditions for S true. In other words, for all formulas $\theta \in F_N$ we have that $T \vdash \forall a S(\theta, a)$ if and only if the Tarski Condition holds for the equivalent formulation of θ in terms of its immediate predecessor.

We now prove the consistency of T via the compactness theorem. Let Γ be a finite set of sentences from T. We thus have that $(M, F_M, S_M) \models \Gamma$ by identifying F with a suitable subset of F_M and S with a suitable restriction S_M. This is clear to see for the case when a formula $\phi \in F_M$. If a formula $\phi \in F_N$ is mentioned in Γ, then since there are only finitely many such formulas in Γ, we can identify it with some formula in $F_M \backslash \Gamma$ so that the requisite Tarski Conditions hold. We therefore have a model $(N, F, S) \models T$.

We can now identify that $M \prec N$ since each $a \in M$ is in T and $(N, F, S) \models Th(M, F_M, S_M)$. We then replace F by F_N by removing any formulas $\psi \in F$ such that $\psi \notin F_N$. We get then, similarly to the above, that $F_M \subseteq F_N$. We then replace S with S_N by the partial satisfaction class obtained from S by only considering formulas $\psi \in F_N$.

Therefore, we have a model $(N, F_N, S_N) \supseteq (M, F_M, S_M)$ as desired. $\square$

It seems highly possible that this method can be adapted so that our new model N is an elementary end extension of M, rather than just an elementary extension. This is a key conjecture for this thesis, which we state below. The reason for the interest in this question is that it allows us to easily construct models with some collection in $\mathscr{L}_S$.

Conjecture 4.1.3. *We can adapt Visser-Enayat's Theorem to obtain an end extension, rather than just an extension. In other words, given* $(M, F_M, S_M) \models \mathrm{PA} + S_M$ *is an* F_M *satisfaction class there exist* N, F_N *and* S_N *such that* $(M, F_M, S_M) \subseteq_e (N, F_N, S_N)$.

We continue the rest of this work modulo an affirmative answer to this question. We assume that, whether through an adaptation of the above theorem or some other means, given $(M, F_M, S_M) \models \mathrm{PA} + S_M$ is an F_M satisfaction class we can construct $(N, F_N, S_N) \models \mathrm{PA} + S_N$ is an F_N satisfaction class such that $(M, F_M, S_M) \subseteq_e (N, F_N, S_N)$.

Theorem 4.1.4. *Suppose* $(N_0, F_0, S_0) \models \mathrm{PA} + S_0$ *is an* F_0 *satisfaction class and given* (N_i, F_i, S_i) *we have that* $(N_{i+1}, F_{i+1}, S_{i+1}) \models \mathrm{PA} + S_{i+1}$ *is an* F_{i+1} *satisfaction class such that* $(N_i, F_i, S_i) \subseteq_e (N_{i+1}, F_{i+1}, S_{i+1})$, *then* $(M, F, S) = \bigcup_{i \in \mathbb{N}} (N_i, F_i, S_i) \models \mathrm{Coll}_{\exists_1}(S)$.

Proof. Let $\theta(x, y, S, \bar{a})$ be an $\exists_1$-$\mathscr{L}_S$-formula and $M \models \forall x{<}t \exists y \theta(x, y, S, \bar{a})$. Without loss of generality we can write the formula $\theta(x, y, S, \bar{a})$ in the form $\exists w \phi(x, y, w, S, \bar{a})$ where ϕ is a quantifier-free $\mathscr{L}_S$-formula. Thus, we have that for some $i \in \mathbb{N}$:

$$N_i \models \forall x{<}t \exists y, w \phi(x, y, w, S_i, \bar{a}).$$

Hence for any $c \in N_{i+1}$ we have that:

$$N_{i+1} \models \forall x{<}t \exists \langle y, w \rangle {<} c \phi(x, y, w, S_{i+1}, \bar{a}).$$

Thus, $N_{i+1} \models \exists z \forall x{<}t \exists \langle y, w \rangle {<} c \phi(x, y, w, S_{i+1}, \bar{a})$ and therefore we conclude that $M \vDash \exists z \forall x{<}t \exists y{<}z \theta(x, y, S, \bar{a})$. $\square$

The reason that this proof is stated for $\exists_1$ formulas is because this ensures that, aside from the first existential quantifier which can be moved 'outside' the formula, the formula is quantifier free. Whilst it may be possible to improve this result, there is difficulty in ensuring that there will always be a suitable S_i which satisfies any quantified subformula we could consider. In $\mathscr{L}_A$ this result can be improved, because we have that the class $\exists_1$ is equal to the class Σ_1. This is known as the MRDP Theorem.

Theorem 4.1.5 (MRDP Theorem)**.** *Within $\mathscr{L}_A$ the class $\exists_1 = \Sigma_1$ [12, Result 7.11].*

If this theorem were to also hold in $\mathscr{L}_S$, then we could improve our theorem to have the strong conclusion that $(M, F, S) \vDash \mathrm{Coll}_{\Sigma_1}(S)$.

Question 4.1.6. *Is there an analogue to the MRDP Theorem in the language $\mathscr{L}_S$? In other words, is it the case that $\exists_1(S) = \Sigma_1(S)$?*

4.2 Ciéśliński's Lemma

In the last section we saw a potential construction to give $\exists_1$ collection in the language $\mathscr{L}_S$ and how this proof may be able to be improved to Σ_1 collection. We will now look at a different way to do this, using a method demonstrated by Ciéśliński which allows translation between Δ_0 sentences in $\mathscr{L}_S$. Using this method, which we call Ciéśliński's method, we shall provide a different approach to give Σ_1 collection in the language $\mathscr{L}_S$.

Lemma 4.2.1 (Císliński's Lemma). *Suppose* $(M, S) \models \Delta_0 - \mathrm{PA}(S)$. *We have that for all standard* Δ_0 *formulas* $\phi(x)$ *in* $\mathscr{L}_S$ $(M, S) \models \forall c[\phi(c) \leftrightarrow S(F_\phi(c'), a)]$ *where* $F_\phi(x) : M \to M$ *is an arithmetically (in* M*) definable* $\mathscr{L}_A$ *sentence and* c' *is the assignment mapping* $x \to c$ *[4, Page 10].*

We have seen however that $\Delta_0 - \mathrm{PA}(S)$ is a very strong base theory to work with, and it would be useful to have a similar lemma available for weaker base theories. We can do this, if we introduce the following two new Tarski Conditions.

Definition 4.2.2. *We introduce the following two new Tarski Conditions:*

Tarski Condition 9) *For all* $\mathscr{L}_A$*-formulas of the form* $\bigwedge_{i<t} \phi_i(x)$:

$$S(\ulcorner \bigwedge_{i<t} \phi_i(x) \urcorner, a) \leftrightarrow \forall i{<}t S(\ulcorner \phi_i(x) \urcorner, a).$$

Tarski Condition 10) *For all* $\mathscr{L}_A$*-formulas of the form* $\bigvee_{i<t} \phi_i(x)$:

$$S(\ulcorner \bigvee_{i<t} \phi_i(x) \urcorner, a) \leftrightarrow \exists i{<}t S(\ulcorner \phi_i(x) \urcorner, a).$$

These Tarski conditions are intuitively acceptable ones to add and are not consequences of the previous Tarski conditions as the following theorem shows. It also tells us that they can be highly useful, as they deal with some of the more problematic pathologous examples that some satisfaction classes make true.

Theorem 4.2.3. *If* $(M, S) \models \mathrm{PA} + S$ *is a full satisfaction class* $+S$ *satisfies Tarski Conditions 9) and 10), then* $(M, S) \models \neg S(\ulcorner \bigvee_{i<t}(0 = 1) \urcorner, a)$ *for any* $t \in M$.

Proof. We know that $(M, S) \models \neg S(\ulcorner 0 = 1 \urcorner, a)$, since $0 = 1$ is an atomic formula. Therefore, $(M, S) \models \neg \exists i{<}t S(\ulcorner 0 = 1 \urcorner, a)$ and thus by Tarski Condition 10) we have that $(M, S) \models \neg S(\ulcorner \bigvee_{i<t}(0 = 1) \urcorner, a)$. □

Thus, this is a highly useful property for our model to posess. It is an open question as to the exact consistency strength of this theory, although we know from Ciéśliński [4, Theorem 4] that this is at most the strength of $\Delta_0 - \mathrm{PA}(S)$.

Question 4.2.4. *What relative consistency strength does the theory* $\mathrm{PA} + S$ *is a full satisfaction class* $+S$ *satisfies Tarski Conditions 9) and 10) have over the theory* $\mathrm{PA} + S$ *is a full satisfaction class?*

The main use of these Tarski Conditions are that they allow us to get the result of Lemma 4.2.1 without, potentially, requiring the full strength of $\Delta_0 - \mathrm{PA}(S)$.

Lemma 4.2.5. *Suppose* $(M, S) \models \mathrm{PA} + S$ *is a full satisfaction class* $+S$ *satisfies Tarski Conditions 9) and 10). We have that for all standard* Δ_0 *formulas* $\phi(x)$ *in* $\mathscr{L}_S$ $(M, S) \models \forall c[\phi(c) \leftrightarrow S(F_\phi(c'), a)]$ *where* $F_\phi(x) : M \to M$ *is an arithmetically (in* M*) definable* $\mathscr{L}_A$ *sentence and* c' *is the assignment mapping* $x \to c$.

Proof. We prove this via induction on the complexity (number of connectives) of $\phi(x)$. We first deal with our base cases. If $\phi(x)$ is a standard $\mathscr{L}_A$ formula, then we define $F_\phi(c')$ as $subst(\phi(x), clterm(c))$. If $\phi(x)$ is $S(\ulcorner\theta\urcorner, x)$, then $F_\phi(c')$ is defined as $subst(\ulcorner\theta\urcorner, c')$.

We now assume that the theorem holds for all formulas with complexity strictly less than ϕ.

If ϕ is of the form $\neg\psi$, then we can take $F_\phi(c') = \ulcorner\neg\urcorner\frown\llcorner F_\phi(c')\lrcorner\urcorner$. Then, $(M, S) \models \phi(c)$ if and only if $(M, S) \models \neg\psi(c)$ which by inductive hypothesis holds if and only if $(M, S) \models \neg S(F_\psi(c'), a)$, which by Tarski Condition 5) holds if and only if $(M, S) \models S(F_\phi(c'), a)$.

If ϕ is of the form $\alpha \wedge \beta$, then we can take $F_\phi(c') = \ulcorner(^\frown \llcorner F_\alpha(m) \lrcorner^\frown \wedge^\frown \llcorner F_\beta(c') \lrcorner^\frown)\urcorner$. Then, $(M, S) \models \phi(c)$ if and only if $(M, S) \models \alpha(c) \wedge \beta(c)$. This holds by induction if and only if $(M, S) \models S(F_\alpha(c'), a) \wedge S(F_\beta(c'), a)$ which is true if and only if $(M, S) \models S(F_\phi(c'), a)$.

If ϕ is of the form $\alpha \vee \beta$, then we can take $F_\phi(c') = \ulcorner(^\frown \llcorner F_\alpha(m) \lrcorner^\frown \vee^\frown \llcorner F_\beta(c') \lrcorner^\frown)\urcorner$. Then, $(M, S) \models \phi(c)$ if and only if $(M, S) \models \alpha(c) \vee \beta(c)$. This holds by induction if and only if $(M, S) \models S(F_\alpha(c'), a) \vee S(F_\beta(c'), a)$ which is true if and only if $(M, S) \models S(F_\phi(c'), a)$.

If ϕ is of the form $\forall x{<}t\theta(x)$, then we can take $F_\phi(c') = \ulcorner \bigwedge^\frown_{b<t} \llcorner F_\theta(c'[^x/_b] \lrcorner \urcorner$. Then we have that $(M, S) \models \phi(c)$ if and only if $(M, S) \models \forall x{<}t\theta(x, c)$. By induction, this holds if and only if $(M, S) \models \forall b{<}tS(F_\theta(c'[^x/_b], a)$. Now, by Tarski Condition 9) we have that this is true if and only if $(M, S) \models S(F_\phi(c'), a)$.

If ϕ is of the form $\exists x{<}t\theta(x)$, then we can take $F_\phi(c') = \ulcorner \bigvee^\frown_{b<t} \llcorner F_\theta(c'[^x/_b] \lrcorner \urcorner$. Then we have that $(M, S) \models \phi(c)$ if and only if $(M, S) \models \exists x{<}t\theta(x, c)$. By induction, this holds if and only if $(M, S) \models \exists b{<}tS(F_\theta(c'[^x/_b], a)$. Now, by Tarski Condition 10) we have that this is true if and only if $(M, S) \models S(F_\phi(c'), a)$. □

We are now able to use this result and Theorem 4.1.4 to produce models of Σ_1 collection in the language $\mathscr{L}_S$.

Theorem 4.2.6. *Suppose* $(N_0, F_0, S_0) \models \mathrm{PA} + S_0$ *is an* F_0 *satisfaction class which satisfies Tarski Conditions 9) and 10). Further, suppose that given*

(N_i, F_i, S_i) we have that $(N_{i+1}, F_{i+1}, S_{i+1}) \models \mathrm{PA} + S_{i+1}$ is an F_{i+1} satisfaction class which satisfies Tarski Conditions 9) and 10) such that $(N_i, F_i, S_i) \subseteq_e (N_{i+1}, F_{i+1}, S_{i+1})$. Then, $(M, F, S) = \bigcup_{i \in \mathbb{N}} (N_i, F_i, S_i) \models \mathrm{Coll}_{\Sigma_1}(S)$.

Proof. Let $\theta(x, y, S, \bar{a})$ be a (standard) Σ_1 sentence in the language $\mathscr{L}_S$ such that $(M, F, S) \models \forall x{<}t \exists y \theta(x, y, S, \bar{a})$. Without loss of generality, we can write $\theta(x, y, S, \bar{a})$ in the form $\exists w \phi(x, y, w, S, \bar{a})$ where $\phi(x, y, w, S, \bar{a})$ is a Δ_0 $\mathscr{L}_S$-sentence. Then using Lemma 4.2.5 we have that:

$$(M, F, S) \models \forall x{<}t \exists y \exists w S(F_\phi(\bar{a}'), \alpha)$$

for a suitable assignment α. Thus, since $\exists w S(f_\phi(\bar{a}'), \alpha)$ is an $\exists_1$ sentence of $\mathscr{L}_S$ we can use Theorem 4.1.4 to get that:

$$(M, F, S) \models \exists z \forall x{<}t \exists y{<}z \exists w S(F_\phi(\bar{a}'), \alpha).$$

Therefore, we can conclude again using Lemma 4.2.5 that:

$$(M, F, S) \models \exists z \forall x{<}t \exists y{<}z \theta(x, y, S, \bar{a}).$$

□

4.3 Modifying Visser-Enayat's Theorem

We see that Tarski Conditions 9) and 10) are desirable properties and thus we want to be able to produce satisfaction classes for a given model where these hold. The natural approach would be to utilise the Visser-Enayat theorem, but we cannot build Tarski Conditions 9) and 10) into a satisfaction class using the Visser-Enayat construction immediately. This is because if we initially add formulas $\theta_1, \theta_2, ..., \theta_n$ to our satisfaction class and some formula $\theta_j = \bigwedge_{i<a} \phi_i$ for some nonstandard a, then we need to add all of these subformulas ϕ_i to our satisfaction class. In the case where each of these formulas is a subformula of another, then this requires infinitely-many steps.

There are various different options to add these axioms to our satisfaction class, we could weaken our Tarski Conditions 9) and 10), strengthen our satisfaction class or restrict the types of formulas that we consider. We shall now consider each of these options in turn to look at their benefits and drawbacks.

4.3.1 Weakening Tarski Conditions 9) and 10)

Our issue in using the Visser-Enayat Lemma to construct satisfaction classes where Tarski Conditions 9) and 10) hold is that a nonstandard-fold disjunction can entail the requirement of infinitely-many steps. Thus, if we only consider disjunctions which contain only finitely many subformulae, there are only finitely-many steps which need completion.

Definition 4.3.1. *We instead consider the following two alternative Tarski Conditions, for some fixed $n \in \mathbb{N}$:*

Tarski Condition $9)^\star_n$ *For all $\mathscr{L}_A$-formulas of the form $\bigwedge_{i<t} \phi_i(x)$, where there are at most n distinct formulas θ_i, we have that:*
$S(\ulcorner \bigwedge_{i<t} \phi_i(x) \urcorner, a) \leftrightarrow \forall i{<}t S(\ulcorner \phi_i(x) \urcorner, a)$.

Tarski Condition $10)^\star_n$ *For all $\mathscr{L}_A$-formulas of the form $\bigvee_{i<t} \phi_i(x)$, where there*

are at most n distinct formulas θ_i, we have that:

$S(\ulcorner \bigvee_{i<t} \phi_i(x) \urcorner, a) \leftrightarrow \exists i{<}t S(\ulcorner \phi_i(x) \urcorner, a)$.

By adding Tarski Conditions $9)^{\star}_{n}$ and $10)^{\star}_{n}$ we avoid the problem of having to complete infinitely-many steps, since only $n \in \mathbb{N}$ formulas need consideration.

Conjecture 4.3.2. *We can construct satisfaction classes using the Visser-Enayat method which satisfy Tarski Conditions $9)^{\star}_{n}$ and $10)^{\star}_{n}$ for each $n \in \mathbb{N}$.*

This weakening of Tarski Conditions 9) and 10) still allows us to deal with the pathologous Examples E.5.1 and E.5.2, but does not provide the full power of Tarski Conditions 9) and 10). In particular, it appears that in a model satisfying these, the conclusion to Lemma 4.2.5 is not necessarily true and thus they cannot be used to obtain $\mathrm{Coll}_{\Sigma_1}(S)$.

4.3.2 Strengthening the Satisfaction Class

We can instead achieve the full power of Tarski Conditions 9) and 10) by allowing the satisfaction class to decide the truth of nonstandard-many formulas simultaneously. If this is the case, our construction no longer requires infinitely many steps, as the nonstandard-many subformulas can all be added simultaneously. We shall call a satisfaction class which can do this a codable satisfaction class.

Definition 4.3.3. *We call a satisfaction class codable if $\{\phi_i : i < \alpha\}$ is an M-finite set, ϕ_i are ${}^{\star}\mathscr{L}_A$ formulas for each i and a_i are a set of assignments, then $\exists t \forall i{<}\alpha[(t)_i \neq 0 \leftrightarrow S(\ulcorner \phi_i \urcorner, a_i)]$.*

If a satisfaction class is codable, then it satisfies Tarski Conditions 9) and 10).

Lemma 4.3.4. *A codable satisfaction class satisfies Tarski Conditions 9) and 10).*

Proof. Let $(M, S) \vDash \mathrm{PA} + S$ is a full codable satisfaction class over M and $\{\phi_i : i < \alpha\}$ is an M-finite set of $^\star\mathscr{L}_A$ formulas. We show this for Tarski Condition 9), the proof for condition 10) is similar.

First suppose that $(M, S) \vDash \forall i{<}\alpha S(\phi_i(x), a)$. Define the formula $\psi_n(x) = \bigwedge_{i<n} \phi_i(x)$. Let t be such that $(M, S) \vDash (t)_n \neq 0 \leftrightarrow S(\ulcorner\psi_n\urcorner, a)$. Then we have that $(t)_0 \neq 0$. We know that if $(M, S) \vDash (t)_n \neq 0$, then $(M, S) \vDash S(\ulcorner\psi_n\urcorner, a)$ which entails that $(M, S) \vDash S(\ulcorner\bigwedge_{i<n} \phi_i(x)$. Then, we have by Tarski condition 6) and 4) that $(M, S) \vDash S(\ulcorner\bigwedge_{i<n+1} \phi_i(x)\urcorner, a)$, so $(M, S) \vDash S(\ulcorner\psi_{n+1}\urcorner, a)$ and therefore $(M, S) \vDash (t)_{n+1} \neq 0$. We can now perform induction on $(t)_n$ to get that $(M, S) \vDash (t)_\alpha \neq 0$ and therefore $(M, S) \vDash S(\ulcorner\bigwedge_{i<alpha} \phi_i(x)\urcorner, a)$ as desired.

Now suppose $(M, S) \vDash S(\ulcorner\bigwedge_{i<\alpha} \phi_i(x)\urcorner, a)$. By repeating the above steps, we see that if $(t)_n = 0$ for some n, then $(t)_\alpha = 0$. We know, however, that $(M, S) \vDash S(\ulcorner\psi_\alpha(x)\urcorner, a)$ and so $(M, S) \vDash t_\alpha \neq 0$. Therefore $(M, S) \vDash \forall i{<}\alpha[(t)_i \neq 0]$ and thus we conclude $(M, S) \vDash \forall i{<}\alpha S(\ulcorner\phi_i(x)\urcorner, a)$. □

The strengthening of a satisfaction class to a codable satisfaction class achieves what we desire, however it appears to be a strong property to add to our satisfaction class with strong unintended consequences. It appears that there will be a similar method to the above proof which shows that $(M, F, S) \vDash \Delta_0 - \mathrm{PA}(S)$, in which case not only does a model with a codable satisfaction class prove $con(\mathrm{PA})$, but also it already has Tarski Conditions 9) and 10) from $\Delta_0 - \mathrm{PA}(S)$.

Conjecture 4.3.5. *The theory of* $\mathrm{PA} + S$ *is a full codable satisfaction class has the same provability strength as* $\Delta_0 - \mathrm{PA}(S)$*.*

Therefore, whilst this is a possible option, we shall instead restrict our class of formulas to only a certain type, to try and reduce the strength of this theory.

4.3.3 Formulas are Stratified

If we restrict our set of formulas F to only consider formulas of finite 'depth', then when adding subformulas of a nonstandard disjunction it is only required to perform finitely-many steps. To explain what we mean by this we first set up the following definition of the rank of a formula and what it means for a formula to be stratified.

Definition 4.3.6. *We define the rank of a formula* θ *in the following inductive manner:*

- $rank(\theta) = 0$ *if* θ *is atomic.*
- $rank(\theta) = 1 + rank(\alpha)$ *if* θ *is of the form* $\neg\alpha$
- $rank(\theta) = 1 + rank(\alpha(x))$ *if* θ *is of the form* $\exists x\alpha(x)$
- $rank(\theta) = 1 + rank(\alpha(x))$ *if* θ *is of the form* $\forall x\alpha(x)$
- $rank(\theta) = 1 + max(rank(\alpha), rank(\beta))$ *is of the form* $\alpha \wedge \beta$

- $rank(\theta) = 1 + max(rank(\alpha), rank(\beta))$ *is of the form* $\alpha \vee \beta$

From this, we can define what it means for a formula to be stratified.

Definition 4.3.7. *We defined a formula* θ *as stratified using the following inductive definition:*

- θ *is stratified, if* θ *is atomic.*
- θ *is stratified, if* θ *is of the form* $\neg\alpha$ *and* α *is stratified.*
- θ *is stratified, if* θ *is of the form* $\exists x \alpha(x)$ *and* $\alpha(x)$ *is stratified.*
- θ *is stratified, if* θ *is of the form* $\forall x \alpha(x)$ *and* $\alpha(x)$ *is stratified.*
- θ *is stratified, if* θ *is of the form* $\alpha \wedge \beta$*,* α *and* β *are stratified and* $rank(\alpha) = rank(\beta)$*.*
- θ *is stratified, if* θ *is of the form* $\alpha \vee \beta$*,* α *and* β *are stratified and* $rank(\alpha) = rank(\beta)$*.*

Therefore, when we look at a stratified formula θ and form a tree of its subformulas, we have that all paths from a leaf to the root have the same (finite) length. Hence, if θ is a conjunction of subformulas, we have that all of these subformulas have the same rank and thus are not subformulas of each other. Therefore, if we restrict F to consider only stratified formulas, then we have that Tarski Conditions 9) and 10) can be built into the Visser-Enayat construction since we do not have infinitely many subformulas to consider in one step.

Theorem 4.3.8. *Suppose we have* $(M, F_M, S_M) \vDash \mathrm{PA} + S_M$ *is an* F_M *satisfaction class such that* F_M *is a set of stratified formulas closed under immediate predecessor. Given a set of stratified* $^\star\mathscr{L}_A$*-formulas* $F_N \supseteq F_M$ *such that we can injectively extend* $\ulcorner \cdot \urcorner$ *so that* $\ulcorner F_N \urcorner \subseteq M$ *and* F_N *is closed under immediate predecessor, there exists* N *and* S_N *such that* $(N, F_N, S_N) \vDash \mathrm{PA} + S_N$ *is an* F_N *satisfaction class* $+ S_N$ *satisfies Tarski conditions 9) and 10) and* $(M, F_M, S_M) \subseteq (N, F_N, S_N)$.

Proof Sketch. We complete the construction in the proof of Theorem 4.1.2, and at the stage when we ensure the Tarski conditions are satisfied for the formulas we also ensure that Tarski Conditions 9) and 10) are satisfied. We can do this, since the formulas are all stratified and thus any nonstandard conjunction of subformulas only requires finitely many steps to be performed. The construction is then completed as in the proof of Theorem 4.1.2. □

This gives us the following corollary that we can generate models of $\mathrm{Coll}_{\Sigma_1}(S)$ with no additional requirements other than we restrict our consideration to stratified formulas.

Corollary 4.3.9. *Suppose* $(N_0, F_0, S_0) \vDash \mathrm{PA} + S_0$ *is an* F_0 *satisfaction class* $+ S_0$ *satisfies Tarski Conditions 9) and 10), where* F_0 *is a set of stratified formulas. Given* $(N_i, F_i, S_i) \vDash \mathrm{PA} + S_i$ *is an* F_i *satisfaction class* $+ S_i$ *satisfies Tarski Conditions 9) and 10), where* F_i *is a set of stratified formulas, we have that* $(N_{i+1}, F_{i+1}, S_{i+1}) \vDash \mathrm{PA} + S_{i+1}$ *is an* F_{i+1} *satisfaction class* $+ S_{i+1}$ *satisfies Tarski Conditions 9) and 10), where* F_{i+1} *is a set of stratified formulas, such that* $(N_i, F_i, S_i) \subseteq_e (N_{i+1}, F_{i+1}, S_{i+1})$*, then* $(M, F, S) = \bigcup_{i \in \mathbb{N}} (N_i, F_i, S_i) \vDash \mathrm{Coll}_{\Sigma_1}(S)$.

Whilst this achieves our aim, it may appear to restrict the amount of formulas we can consider widely. Clearly there are many formulas of interest which are not stratified. It is the case, however, that every formula can be written in an equivalent stratified form.

Lemma 4.3.10. *For any $\mathscr{L}_A$-formula θ, there exists a stratified $\mathscr{L}_A$-formula S_θ such that $(M, S) \models \theta \leftrightarrow S_\theta$.*

Proof. We prove this via induction on the complexity (number of connectives) of θ. If θ is complexity 0, then it is atomic and therefore balanced. We now assume it holds for all formulas of complexity strictly lower than θ. If our formula is $\neg\theta$, then by our hypothesis θ is equivalent to some balanced formula η and thus $\neg\theta$ is equivalent to the balanced formula $\neg\eta$. The case for $\exists x\theta(x)$ and $\forall x\theta(x)$ is similar. If our formula is $\alpha \wedge \beta$, then by our inductive hypothesis α is equivalent to a balanced formula ϕ and β is equivalent to a balanced formula ψ. Suppose that $\text{rank}(\phi) = \text{rank}(\psi) + n$, then we have that the formula $((...((\phi\wedge\phi)\wedge(\phi\wedge\phi))\wedge((\phi\wedge\phi)\wedge(\phi\wedge\phi))...))\wedge\psi$ is balanced and equivalent to $\alpha \wedge \beta$, where there are 2^n instances of ϕ. The case for $\alpha \vee \beta$ is proven similarly. □

Thus, we can, indirectly, consider the truth of any formula and build a satisfaction class which satisfies $\text{Coll}_{\Sigma_1}(S)$. However, it is not clear whether this is a true partial satisfaction class or not and whether it satisfies Lachlan's theorem. If not, it could be the case that every model of PA, M, has satisfaction class for stratified formulas. This is a highly interesting open question, which has interesting ramifications for a deflationist theory of truth.

Question 4.3.11. *Does a structure (M, F, S) where $M \models \mathrm{PA}$ and S is an F-satisfaction class and F is a set of stratified formulas satisfy Lachlan's Thorem?*

Question 4.3.12. *Given M, F where $M \models \mathrm{PA}$ and F is a set of stratified formulas, does there exist S such that S is an F-satisfaction class?*

Chapter 5

Deflationary Truth and Satisfaction Classes

In this chapter we provide a formalised definition of what it means for a truth predicate to be deflationary and explore the consequences that this for has the deflationist position.

Ketland, as seen in Section C.4, provides the following definition of a deflationary truth predicate.

Definition 5.1 (Ketland)**.** *The truth theory T with predicate* Tr^M *is deflationary if for all models* $(M, \mathrm{Tr}^M) \vDash \mathrm{PA} + T$ *we have that* $(M, \mathrm{Tr}^M) \vDash \mathrm{Tr}^M(\ulcorner\sigma\urcorner) \leftrightarrow \sigma$ *for all sentences σ in the signature of M. Further, we require that T is conservative over* PA. *[13, Pages 71–72]*

Whilst this definition expresses that a sentence of the model is true if and only if its meaning holds in the model, it perhaps does not fully adequately capture the more vague notion that truth has no metaphysical nature beyond the first property. There exist conservative theories T over PA such that we cannot always expand any model $M \models \text{PA}$ so that $(M, ...) \models T$. Therefore, whilst the theory T is conservative, it still appears to have a facet that is inaccessible to M. Therefore, we shall adapt this definition to provide our own definition with the more stringent requirement that every model M of interest must have a unique expansion such that $(M, \text{Tr}^M) \models T$. Instead, we provide our own loose definition, to be clarified later, below.

Definition 5.2. *The truth theory T with predicate Tr^M is deflationary if for all models $(M, \text{Tr}^M) \models \text{PA}+T$ we have that $(M, \text{Tr}^M) \models \text{Tr}^M(\ulcorner\sigma\urcorner) \leftrightarrow \sigma$ for all sentences σ of consideration. Further, T is only deflationary if every model $M \in \mathscr{C}$ where $\mathscr{C}$ is the considered class of models of* PA *has an expansion such that $(M, \text{Tr}^M) \models T$.*

In this definition we state that for a truth predicate to be deflationary it must both satisfy the equivalence schema, that $(M, \text{Tr}^M) \models \sigma \leftrightarrow \text{Tr}^M(\sigma)$ for all considered sentences σ and every model $M \in \mathscr{C}$ must have an expansion to (M, Tr^M), for some class of models of PA, $\mathscr{C}$. The equivalence schema is clearly required for a truth predicate to be deflationary, but whether every model requires an expansion to a truth predicate is less obvious. Some property is required, to capture the notion that deflationary truth has no metaphysical content, but this criterion differs to Ketland's, in that he requires only syntactic conservativity.

The reason for taking this semantic, rather than syntactic approach, is that a deflationary truth predicate has a semantic interpretation. If we consider truth predicates syntactically, as per Ketland, then given a theory T we are building a new theory $T' = T$ + a new predicate f + axioms for how f

behaves. For the theory $T' \vdash \sigma \leftrightarrow f(\sigma)$ to have any utility we require some interpretation of what f is. Under this examination, without any semantic interpretation of f, f appears like an alternative definition of the identity function.

The semantic interpretation of f is that this corresponds to a notion of truth. This notion of truth cannot be model-truth, since that is what we are defining, and therefore must be meta-truth, as per the definitions of model-truth and meta-truth in Section C.3. Thus, $T' \vdash f(\sigma)$ has the intended interpretation that $\mathrm{Tr}^m(T \vdash \sigma)$. From the soundness and completeness of first order logic, however, this means that $\mathrm{Tr}^m(T \vDash \sigma)$, in other words $\mathrm{Tr}^m(M \vDash \sigma$ for all $M \vDash T)$. Applying the Tarski Axioms, or alternatively simple intuition on the behaviour of truth, to this means that $\mathrm{Tr}^m(M \vDash \sigma)$ for all $M \vDash T$. Therefore, under the intended interpretation, $M \vDash f(\sigma)$ for all $M \vDash T$. It is clear to see, therefore, that for f to behave as a desired truth-predicate, every model must have an expansion to such an f. Therefore, we find Ketland's definition is lacking and syntactic conservativity is not enough.

It is important to note that, without the desire for truth to have no metaphysical content, the above argument is not valid. For example, suppose truth has the substantival facet of corresponding to formal consistency. In which case, from $\mathrm{Tr}^m(T \vdash \sigma)$ we instead derive that $\mathrm{Tr}^m(M \vDash \sigma$ for all $M \vDash T + con(T) + con(con(T)) + ...)$. The reason for this is because the truth predicate corresponding to formal consistency states that anything other models believe is false.

The conclusion to this is that the definition provided in Definition 5.2 gives at least necessary conditions for a truth predicate to be considered deflationary. Whether this definition is sufficient, however, is not quite so obvious. It could be argued that every model not only requires an expansion to a truth predicate, but to a unique truth predicate. An example of a truth predicate which doesn't satisfy this, is the notion of a satisfaction class as provided in Definition E.2.3. If the sentences under consideration are the standard $\mathscr{L}_A$ sentences, then by Theorem E.2.6 every satisfaction class satisfies the equivalence schema. By Theorem E.2.8, however, a countable recursively saturated model of PA has $2^{\aleph_0}$ different satisfaction classes.

The question is raised by a model having many distinct truth predicates as to which one of these is the 'correct' truth predicate. Intuitively, there is only one meta-truth predicate. Therefore, by analogy, there should only be one model-truth predicate also. If a model has many such truth-predicates, then choosing which one of these to use requires some judgement which is not based on the theory, meaning truth has this additional nature and is not deflationary. In the case with satisfaction classes, this argument is flawed in that it assumes we need a judgement on which satisfaction class is the truth predicate. One can take an arbitrary satisfaction class and if one is only concerned with $\mathscr{L}_A$ sentences, then any satisfaction class will suffice. By analogy, if one is only interested in whether PA $\vDash \forall x \exists y [y = x + 1]$, then one considers an arbitrary model $M \vDash$ PA and does not make a judgement on which model is the 'correct' model.

Therefore, we shall take Definition 5.2 as a suitable characterisation of a deflationary truth predicate. Given this, examples of truth predicates which aren't deflationary are easy to find. By Lachlan's theorem we know that only recursively saturated models have a satisfaction class. Therefore, if the class of models under consideration includes models of PA which aren't recursively saturated, then a satisfaction class is not a deflationary truth predicate. Further, the theory of $\Delta_0 - \text{PA}(S)$ is strong enough to prove the consistency of PA. Thus, if the class of models under consideration includes any model $M \vDash \neg con(\text{PA})$, then a truth predicate from this theory is also not deflationary.

We therefore arrive at the question, as to how to choose between these truth predicates. The answer to this, clearly depends on an answer to the question of which sentences and which models to consider. It is an immediate consequence of Tarski's Theorem, Theorem E.1.10, that the language cannot be the entirety of $\mathscr{L}_{Tr}$, where this is the language of the theory with a truth predicate. This is because we can define a sentence σ in this language such that $(M, Tr) \vDash \sigma \leftrightarrow \neg Tr(\ulcorner \sigma \urcorner, a)$ which contradicts the equivalence schema.

As a minimum, the sentences under consideration should belong to the language $\mathscr{L}_A$. It is in this language that we express and discuss the truths of arithmetic and is the background language that we have worked within. More controversial, is whether the sentences that we should work with are all necessarily standard. In the real world it appears that we work with only sentences of finite length and only speak about the truth and falsity of these sentences. This, however, on closer examination is not the case. For example, take $n \in \mathbb{N}$, then the sentence $\text{Tr}^m(\mathbb{N} \vDash \forall x{<}\dot{n}[x \geqslant 0]$. is a natural assertion. The sentence $\forall x{<}\dot{n}[x \geqslant 0]$ is a standard sentence of $\mathscr{L}_A$, but can be rewritten as the sentence $0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{n} - 1) \geqslant 0$, which is also a standard sentence of $\mathscr{L}_A$. Similarly, take a nonstandard model $M \vDash \text{PA}$ and a nonstandard element $a \in M$. Then the sentence $\text{Tr}^m(M \vDash \forall x{<}\dot{a}[x \geqslant 0])$ is also one that we desire to assert. Similarly, the sentence $\forall x{<}\dot{a}[x \geqslant 0]$ can be written as the sentence $0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{a}-1) \geqslant 0$. This sentence, however, is a nonstandard sentence of arithmetic, as it has nonstandard length. The assertions:

$$\mathrm{Tr}^m(\mathbb{N} \vDash 0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{n}-1) \geqslant 0) \text{ and}$$

$$\mathrm{Tr}^m(M \vDash 0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{a}-1) \geqslant 0)$$

are entirely similar in structure and, assuming that we can talk about the truth of nonstandard models, both are valid assertions. Given the assertion $\mathrm{Tr}^m(M \vDash 0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{a}-1) \geqslant 0)$ our truth predicate should, therefore, be such that $(M, \mathrm{Tr}^M) \vDash \mathrm{Tr}^M(\ulcorner 0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{a}-1) \geqslant 0 \urcorner)$. This means that the formalised truth predicate must also be able to discuss the truth of nonstandard sentences of arithmetic.

One objection to this, could be that the nonstandard sentence:

$$0 \geqslant 0 \wedge 1 \geqslant 0 \wedge ... \wedge (\dot{a}-1) \geqslant 0$$

can be written as a standard $\mathscr{L}_A$ sentence, so therefore to discuss the truth of these sentences we do not need to allow nonstandard sentences. This objection does not work, however. There are many nonstandard sentences which cannot be translated into a standard sentence. For an example of this, consider the sentence $\exists x_1 \exists x_2 ... \exists x_a \bigwedge_{1 \leqslant i \neq j \leqslant a} x_i \neq x_j$ where a is a nonstandard number. Therefore, at least some of the sentences under consideration shall be nonstandard.

This analysis relied on the assumption that we can talk about the truth of nonstandard models. Given that normal arithmetic is usually only within the standard model, why should this be the case? The theory PA is neutral on whether a model is the standard model or a nonstandard model. Working solely from the theory PA, there is no way of drawing this distinction. Therefore, if truth has no metaphysical content, there should be no way of drawing this distinction using a deflationary model-truth predicate. Further, the meta-truth predicate should apply to nonstandard models as well. In the same way we can talk about the truth of sentences within $\mathbb{N}$, $\mathbb{R}$, graphs, groups and many other mathematical structures, there is no reason not to talk of truth within nonstandard models as well. Lastly, if one does want to exclude nonstandard models, then this can be done by considering the second order theory of induction with PA^-, PA^2 which has only one model, the standard model. This theory, however, is not syntactically complete and thus studying its consequences solely from syntactic derivability is not possible. Because of this, we instead have looked at the first order theory PA which is both sound and complete.

Therefore, the class of models under consideration shall at least include some nonstandard models of arithmetic. It might be tempting to restrict the class of models to only those models which believe that PA is consistent. This would allow the deflationist to regard the very strong theory $\text{PA}(S)$ as a formal deflationary theory of truth. Further support for this, comes from the fact that we believe that PA is consistent. However, similarly to nonstandard models, the theory PA does not draw a distinction between models of $con(\text{PA})$ and $\neg con(\text{PA})$ from Gödel's Second Incompleteness Theorem.

It may similarly be tempting to restrict the class of models to only those models which are recursively saturated, so that the deflationist can regard satisfaction classes as a formal deflationary theory of truth. This is a weaker condition than the other two restrictions, since every model has an elementary extension which is recursively saturated by Theorem A.2.10. Again, however, the theory PA has no formal distinction between models which are recursively saturated and those which are not.

We propose that the models under consideration should be every model of the theory. Colloquially, we talk of a sentence being true inside a certain model, for any model of a theory. Indeed, as observed above, we use meta-truth to write $\mathrm{Tr}^m(\mathrm{PA} \vDash \sigma)$. For the deflationist, this must translate to $\mathrm{Tr}^m(M \vDash \sigma \text{ for all } M \vDash \mathrm{PA})$, which translates to $\mathrm{Tr}^m(M \vDash \sigma)$ for all $M \vDash \mathrm{PA}$. This tells us that meta-truth applies to every model of PA and so for a truth predicate to be deflationary, every model must have an expansion to one.

Finally, we arrive at the following definition of a deflationary truth predicate.

Definition 5.3. *The truth theory T with predicate Tr^M is deflationary if for all models $(M, \mathrm{Tr}^M) \vDash \mathrm{PA} + T$:*

$$(M, \mathrm{Tr}^M) \vDash \mathrm{Tr}^M(\ulcorner \sigma \urcorner) \leftrightarrow \sigma$$

for all nonstandard sentences σ. Further, T is only deflationary if every model $M \vDash \mathrm{PA}$ has an expansion such that $(M, \mathrm{Tr}^M) \vDash T$.

This definition has some interesting consequences for the theories of satisfaction classes in Section E.5. The theory of $\mathrm{PA}(S)$ is a desirable theory of model-truth, but entails the consistency of PA by Theorem E.5.5 and thus, by our definition, is not a deflationary theory of truth. Similarly, since $\Delta_0 - \mathrm{PA}(S)$ entails the consistency of PA by Theorem E.5.7 this is also not a deflationary theory of truth.

Full satisfaction classes similarly are not a deflationary theory of model truth, since they entail that the model possessing one is recursively saturated, by Lachlan's Theorem, Theorem E.2.9. However, the theory of full satisfaction classes also fails this test for another reason, it does not satisfy the equivalence schema. As shown in Example E.5.1 there exist sentences σ such that $(M, S) \vDash \sigma \wedge \neg S(\ulcorner\sigma\urcorner, a)$. Worse than failing the equivalence schema, however, satisfaction classes also fail what I shall call the weak equivalence schema. This is the schema that for all $\mathscr{L}_A(M)$ sentences σ: $(M, S) \vDash \neg[\sigma \wedge \neg \mathrm{Tr}^M(\ulcorner\sigma\urcorner)]$. This schema is intuitively highly acceptable and surely should hold for our theory of truth.

This leads to the conclusion that the theory of full satisfaction classes is not a satisfactory theory of truth. As it fails the weak equivalence schema, which it appears every theory of truth should satisfy. However, the theory appears to contain the minimum conditions that an axiomatic theory of truth can satisfy for all nonstard sentences of $\mathscr{L}_A$. The Tarski conditions mimic intuition on deflationary model-truth and any formal theory of truth should satisfy these. For example, suppose that $(M, \mathrm{Tr}^M) \vDash \forall x \mathrm{Tr}^M(\ulcorner\theta(x)\urcorner)$. This holds if and only if $\mathrm{Tr}^m(M \vDash \forall x\theta(x))$ which holds if and only if $(M, \mathrm{Tr}^M) \vDash \mathrm{Tr}^M(\ulcorner\forall x\theta(x)\urcorner)$. Since the theory of full satisfaction classes is strong, in that it restricts the number of models we can talk about the truth of, and weak in that it fails the equivalence schema, we arrive at the following theorem.

Theorem 5.4. *There does not exist a model-truth predicate* Tr^M *such that* $(M, \mathrm{Tr}^M) \models \sigma \leftrightarrow \mathrm{Tr}^M(\ulcorner \sigma \urcorner)$ *for all* ${}^*\mathscr{L}_A(M)$*-sentences* σ *such that all models* $M \models \mathrm{PA}$ *have an expansion to* Tr^M*.*

This tells us that if meta-truth is deflationary, then it cannot be formalised for PA for all nonstandard sentences σ, under the given definition. This appears to be very unappealing to the deflationist. However, I do not believe that this is devastating to the deflationist. The deflationist could consider truth for a weakened version of the language, for example the language of stratified formulas as seen in Definition 4.3.7, where a satisfaction class over these may not obey Lachlan's Theorem. Further, for such languages, satisfaction classes may be able to be conservatively strengthened, for example Theorem 4.3.9 shows how Σ_1 collection axioms for the language $\mathscr{L}_S$ can be obtained. This is speculative, but leads to the following open question which is of great interest to the deflationist.

Question 5.5. *Can the nonstandard language of* $\mathscr{L}_A$ *be restricted in such a way so that it loses no expressive content and so that there exists a deflationary truth predicate for* PA *over all sentences of this restricted language?*

This would rescue the deflationist from the plight of not having a deflationary truth predicate for all nonstandard sentences of arithmetic. Further, it would provide justification for choosing this deflationary truth predicate, because it appears to be able to be strengthened considerably. In summary, whilst a good deflationary truth predicate for Peano Arithmetic currently has no known formalisations, it is possible that research into other languages, such as the stratified language, will generate one, which would be of great value to the deflationist.

Chapter 6

Conclusion

In this thesis we set out to provide new formalisations of truth for Peano Arithmetic and to explore the mathematical and philosophical consequences of this.

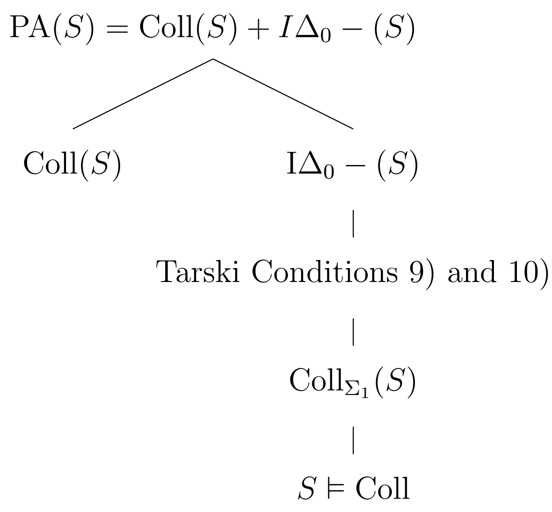
We chose to explore satisfaction classes as the base theory of truth to work from. The reason for this is that satisfaction classes talk about the truth and falseness of all standard and nonstandard sentences of arithmetic. This is the natural base theory to work from when looking at the the truth of nonstandard sentences, because it consists of only six conditions, all of which are philosophically highly acceptable. The theory is therefore open to modification, because it consists of so few conditions. Further, modification is desirable, since satisfaction classes can ensure that intuitively false sentences can be true and intuitively true sentences can be false.

In Chapter 2 we explored our first theory of truth for Peano Arithmetic, that of satisfaction classes which believe all collection axioms are true. We spent the bulk of this chapter proving that this theory is consistent for nonstandard, recursively saturated models of Peano Arithmetic. Given that the theory is consistent, it is a desirable theory of truth, since the collection axioms are believed to be true and are useful axioms to have. This theory, further, is conservative over the base theory of satisfaction classes, by the consistency result. For this reason, we argue that this theory is a better theory of truth, than the base theory, since it is closer towards our intuitive notion of truth, whilst containing no additional unattractive consequences. We then saw at the end, however, a hint that this theory behaves in unexpected ways. Usually, the collection axioms together with induction for Δ_0-formulas is enough to prove induction for all formulas, but the standard proof of this does not work in the context of within satisfaction classes. It is uncertain as to the extent of this observation, but we suggest that this is an interesting area of research for further study.

Chapter 3 begins with an exploration of a stronger theory of collection for satisfaction classes, the theory where all standard collection axioms are true for formulas involving the satisfaction predicate. We show that this theory is at least as powerful as the theory considered above and then note that this theory is also consistent, since it is entailed by the theory of satisfaction classes together with induction axioms for all formulas involving the satisfaction predicate. We also note that there are theories of satisfaction classes which do not have collection axioms for all formulas in the language with the satisfaction predicate. The chapter then produces a definition of a closure for models with a satisfaction class, to produce a submodel with an induced satisfaction class. We conjecture that this method will provide models with a satisfaction class without all collection axioms for the satisfaction predicate, but this method is also hoped to be more generally useful in other circumstances. We finally consider ways of arriving at this stronger theory of collection for satisfaction classes using existing results in the literature and see that they provide easy methods for uncountable models, but not for countable models.

In Chapter 4 we generalise and rephrase Visser and Enayat's construction of end extensions of models with a satisfaction classes to many-sorted logics. We then work, modulo the conjecture that this can be modified to provide elementary end extensions, and show that this allows us to construct models with collection for $\exists_1$-formulas in the language with the satisfaction predicate. We then show that, with two additional Tarski axioms we can simulate Ciéslińksi's method and improve this to models with collection for Σ_1-formulas in the language with the satisfaction predicate. We then consider three methods of building a satisfaction class with these additional Tarski axioms. The first, weakening the axioms, does not provide their full power whilst the second, building a satisfaction class which can consider nonstandardly-many sentences at the same, appears to have the strength of Δ_0-PA(S). Finally, we consider reducing the language to one where every formula has a balanced tree of subformulas. Modulo the conjecture above, satisfaction classes with collection axioms for Σ_1-formulas in the language with the satisfaction class can be built, but interestingly such satisfaction classes may not satisfy Lachlan's theorem.

These chapters provide a variety of different theories of satisfaction class which, in conjunction with the theories within the literature in Subsection E.5, gives rise to the following hierarchical picture.



Finally, in Chapter 5 we argue for a specific definition of what it means for a truth predicate for Peano Arithmetic to be a deflationary truth predicate. We arrive at this conclusion using philosophical argumentation and show that this entails that there is no deflationary truth predicate, in the sense of this definition, for Peano Arithmetic. We then argue, however, that it may be possible for a deflationary truth predicate to exist for a reduced language of arithmetic, perhaps the stratified language as considered in Chapter 4.

Within this thesis we have provided new formalisations of truth for Peano Arithmetic, ones which are attractive and interesting, but also can be hard to conservatively obtain. We have then reflected philosophically, and seen that this has ramifications for a deflationary conception of truth and suggests that it cannot be formalised for the full nonstandard language of arithmetic.

Appendix A

Technical Background

A.1 Peano Arithmetic

We shall work predominantly in the language of arithmetic, $\mathscr{L}_A$. Some results shall be stated in terms of a finite extension of this language, $\mathscr{L}$. A particular special case of this is $\mathscr{L}_S$ which is the language $\mathscr{L}_A \cup \{S\}$, where S is the satisfaction relation. We shall also consider the languages ${}^\star\mathscr{L}_A$ and ${}^*\mathscr{L}_A(M)$, which are defined by Robinson [17, Section 3]. The former allows sentences of nonstandard length and the latter allows sentences of nonstandard length including parameters from a model M.

All technical work shall be completed in the background theory of PA^-, also known as the theory of non-negative parts of discretely ordered rings. For the details of this theory, the reader is referred to Chapter 2 of *Models of Peano Arithmetic* [12]. The majority of results shall be stated within the theory of first order Peano Arithmetic, PA, which is the theory of PA^- together with the axiom schema of induction. Sometimes, we shall work in a fragment of this theory, $PA^- + I\Gamma$, for some class of formulas Γ, where this denotes the theory of PA^- together with induction for all formulas in Γ. Working within PA gives us access to Skolem functions. For the full details of these and PA, the reader is referred to Chapters 4 and 6 of *Models of Peano Arithmetic* [12].

There are also second order variants of Peano Arithmetic. Often, what is meant by this is a first order structure where certain sets of a model can be formulated. In this work, in reference to second order logic, we willl only consider 'true' second order logic, where we may quantify over all subsets of the model.

A.1.1 Important Theorems

Theorem A.1.1 (MacDowell–Specker). *If $M \models \mathrm{PA}$, then there exists $N \models \mathrm{PA}$ such that $M \prec_e N$. [12, Theorem 8.1]*

Theorem A.1.2 (Löwenheim–Skolem Theorems). *If M is an infinite model of a countable first order theory T, then for every infinite cardinal number κ there exists $N \models T$ where $|N| = \kappa$ such that:*

- *If $\kappa < |M|$, then $N \prec_e M$*
- *If $\kappa > |M|$, then $M \prec_e N$ [12, Theorem 0.4 and 0.5]*

Theorem A.1.3 (Tarski–Vaught Test). *If M, N are $\mathscr{L}$-structures such that $M \subseteq N$, then $M \prec N$ if and only if for all $\bar{a} \in M$ and all $\mathscr{L}$-formulas $\theta(x, \bar{a})$ such that $N \models \exists x\theta(x, \bar{a})$ there exists $y \in M$ such that $N \models \theta(y, \bar{a})$ [12, Theorem 0.6].*

Theorem A.1.4 (Gödel–Rosser Incompleteness Theorem). *For any recursive, consistent theory* $T \supseteq$ PA*, there exists an* $\mathscr{L}_A$*-sentence* σ *such that* $T \nvdash \sigma$ *and* $T \nvdash \neg\sigma$ *[12, Corollary 3.15].*

Theorem A.1.5 (Gödel's Second Incompleteness Theorem). *For any recursive theory* $T \supseteq$ PA*, if* $T \vdash con(T)$*, then* T *is inconsistent [12, Exercise 3.11].*

Theorem A.1.6 (Tennenbaum's Theorem). *There is no countable, nonstandard model* $M \models$ PA *such that* M *is recursive [12, Theorem 11.11].*

Theorem A.1.7 (Omitting Types). *Let* Σ *be a consistent set of* $\mathscr{L}$*-sentences and suppose that* p_n *is a set of* $\mathscr{L}$*-formulas* $\phi(x_1, ..., x_{k_n})$ *for each* $n \in \mathbb{N}$*. If for each* $n \in \mathbb{N}$ *there is a formula* $\psi(x_1, \ldots, x_{k_n})$ *such that* $\Sigma \cup \{\exists x \psi(x_1, \ldots, x_{k_n})\}$ *is consistent and* $\Sigma \vdash \forall x[\psi(x_1, \ldots, x_{k_n}) \to \phi(x_1, \ldots, x_{k_n})]$*, then there is a model* $M \models \Sigma$ *where* M *omits* p_n *for each* $n \in \mathbb{N}$*. This means that for all constants* $c_1, \ldots, c_{k_n}$ *there is a formula* $\phi(x_1, \ldots, x_{k_n})$ *in* p *such that* $M \models \neg\phi(c_1, ..., c_{k_n})$ *[11, Theorem 10.36].*

A.1.2 Coding

In the theory PA, there exists the Gödel β Function which allows us to code any finite sequence of numbers as a single number. In particular, given a sequence $(a_0, a_1, ..., a_n)$ we denote the number that codes this sequence by a and have notationally that $(a)_i = \beta(a, i) = a_i$. The existence of this function can be proven with the Chinese Remainder Theorem, the details of which can be found in Chapter 3 of *Models of Peano Arithmetic* [12]. When coding two numbers a single number we use the pairing function, which we denote as $\langle x, y\rangle$ for all numbers x, y.

This allows us to code subsets of models using numbers as well. We do this using the following definition.

Definition A.1.8 (Set coding)**.** *We say that a set $S \subseteq M$ is coded in a model $M \models \mathrm{PA}^-$ if there is some $s \in M$ such that $S = \{i \in M : p_i | s\}$, where p_i is the ith prime number in M. Alternatively, we can write this as $i \in S$ if and only if $p_i | s$. We call such a set an M-finite set [12, Definition 3.10].*

This gives rise to the following useful theorem.

Theorem A.1.9. *A finite set $S \subseteq \mathbb{N}$ is coded in a model M if and only if S is coded by a standard natural number $s \in \mathbb{N}$.*

We are able to code formulas of finitely many variables in a finite language $\mathscr{L}$ using a Gödel-numbering. This is an injective mapping $\ulcorner \cdot \urcorner : \{\theta(\bar{x}) : \theta(\bar{x}) \text{ is an } \mathscr{L}_A - \text{formula}\} \to \mathbb{N}$. The details on the exact definition of a Gödel-numbering are technical and can be found in Chapter 9 of *Models of Peano Arithmetic* [12].

A.1.3 Classes of Sentences

We have various different classes of formulas for a finite language $\mathscr{L} \supseteq \mathscr{L}_A$. We provide definitions of these below.

Definition A.1.10 (Δ_0)**.** *The class Δ_0 is the class of all formulas $\theta(\bar{x})$ such that $\theta(\bar{x})$ contains no unbounded quantifiers [12, Definition 2.9].*

Definition A.1.11 ($\exists_1$ and $\forall_1$)**.** *The class $\exists_1$ is the class of all formulas $\theta(\bar{x})$ such that $\theta(\bar{x})$ is of the form $\exists\bar{y}\phi(\bar{y},\bar{x})$ where $\phi(\bar{y},\bar{x})$ is a quantifier-free formula. The class $\forall_1$ is the class of all formulas $\theta(\bar{x})$ such that $\theta(\bar{x})$ is of the form $\forall\bar{y}\phi(\bar{y},\bar{x})$ where $\phi(\bar{y},\bar{x})$ is a quantifier-free formula [12, Page 99].*

Definition A.1.12 (Σ_n, Π_n and Δ_n)**.** *Let $n \geqslant 1$. By Σ_0 and Π_0 we mean the class Δ_0. The class Σ_n is the class of all formulas $\theta(\bar{x})$ such that $\theta(\bar{x})$ is of the form $\exists\bar{y}\phi(\bar{y},\bar{x})$ where $\phi(\bar{y},\bar{x})$ is a Π_{n-1} formula. The class Π_n is the class of all formulas $\theta(\bar{x})$ such that $\theta(\bar{x})$ is of the form $\forall\bar{y}\phi(\bar{y},\bar{x})$ where $\phi(\bar{y},\bar{x})$ is a Σ_{n-1} formula. The class Δ_n is the class of all formulas $\theta(\bar{x})$ in $\mathscr{L}_A$ such that $\theta(\bar{x})$ is equivalent to a formula that is Π_n and also equivalent to a formula that is Σ_n [12, Definition 7.1].*

Definition A.1.13 (Δ_0^+, Σ_n^+ and Π_n^+)**.** *The class Δ_0^+ is Δ_0. By Σ_0^+ and Π_0^+ we denote the class Δ_0^+. Let $n \geqslant 1$. The class Σ_n^+ is the closure of Π_{n-1}^+ under the connectives $\wedge, \vee$, bounded universal quantification and existential quantification. The class Π_n^+ is the closure of Σ_{n-1}^+ under the connectives $\wedge, \vee$, bounded existential quantification and universal quantification [12, Exercise 7.2]*

A.2 Model Theoretic Notions

We say that the standard model of Peano Arithmetic is the familiar structure of the natural numbers $\mathbb{N}$ with constants $0, 1$ and relations $+, \cdot$ and $<$ interpreted in the usual way. PA also gives rise to nonstandard models of arithmetic, which are any model $M \vDash \mathrm{PA}$ such that $M \ncong \mathbb{N}$. The existence of these models is an easy corollary of the compactness theorem or the Löwenheim-Skolem theorems.

The following example provides a way of generating a new model of PA^- from a former model.

Example A.1. *If $M \models \mathrm{PA}^-$ then there is a model $M[X]^+ \models \mathrm{PA}^-$ such that $M \subseteq_e M[X]^+$. We form the model $M[X]^+ \models \mathrm{PA}^-$ by constructing the discretely ordered ring $M[X]$ of M and then defining an order upon this ring, where $X > a$ for all $a \in M$. We then see that the non-negative part of this ring $M[X]^+$ is a model of PA^- [12, Example 2.3].*

A.2.1 Relations between Models

We have various relations that can hold between models, which we define below.

Definition A.2.1 (Initial Segment and End-extension)**.** *Let M, N be models of a finite language $\mathscr{L}$. We write that $M \subseteq_e N$ and say that M is an initial-segment of N if $M \subseteq N$ and for all $x \in M$ and all $y \in N$ we have that if $N \models y < x$, then $y \in M$. If $M \subseteq_e N$, then we say that N is an end-extension of M.*

We regard $\mathbb{N} \subseteq_e M$ for all models $M \models \mathrm{PA}^-$ [12, Section 2.2].

Definition A.2.2 (Elementary)**.** *Let M, N be models of a finite language $\mathscr{L}$. We write that $M \prec N$ and say that N is an elementary extension of M if $M \subseteq N$ and for all $\mathscr{L}$-formulas $\theta(\bar{x})$ and all $\bar{a} \in M$ we have that $M \models \theta(\bar{a})$ if and only if $N \models \theta(\bar{a})$.*

Let Γ be a class of $\mathscr{L}$-formulas. We say that N is a Γ-elementary extension of M, notationally $M \prec_\Gamma N$ if $M \subseteq N$ and for all $\mathscr{L}$-formulas $\theta(\bar{x}) \in \Gamma$ and $\bar{a} \in M$ we have that $M \models \theta(\bar{a})$ if and only if $N \models \theta(\bar{a})$.

We say that N is an elementary end extension of M, notationally $M \prec_e N$ if $M \subseteq_e N$ and $M \prec N$ [12, Definition 2.10].

We have the following theorem which relates end-extensions and elementarity.

Theorem A.2.3. *If $M \models \mathrm{PA}^-$ and $N \models \mathrm{PA}^-$, then $M \subseteq_e N$ implies that $M \prec_{\Delta_0} N$ [12, Theorem 2.11].*

A useful way of generating elementary submodels, is to consider the closure of a model.

Definition A.2.4. *Let $M \models \mathrm{PA}$. We define $cl_M(A) = \{x \in M :$ for some $\bar{a} \in A$ there is $\theta(x, \bar{a})$ in $\mathscr{L}_A$ such that $M \models \exists x\theta(x, \bar{a}) \wedge \forall y{<}x \neg\theta(y, \bar{a})\}$. If Γ is a class of formulas, then we define $\Gamma - cl_M(A) = \{x \in M :$ for some $\bar{a} \in A$ there is an $\mathscr{L}_A$-formula $\theta(x, \bar{a})$ in Γ such that $M \models \exists x\theta(x, \bar{a}) \wedge \forall y{<}x \neg\theta(y, \bar{a})\}$ [12, Section 6.3].*

The standard closure of a model has the following useful theorem.

Theorem A.2.5. *Let $M \models \mathrm{PA}$. Then $A \subseteq cl_M(A)$ and $cl_M(A) \prec M$ [12, Theorem 6.9].*

A.2.2 Cardinalities of Models

We introduce the concept of a regular cardinal and what it means for model to be κ-like, for some cardinal κ.

Definition A.2.6 (Regular cardinal)**.** *We call a cardinal κ regular if for any set X such that $|X| = \kappa$, then if $X = \bigcup_{i \in I} Y_i$, then $|I| = \kappa$ or $|Y_j| = \kappa$ for some $j \in I$. A cardinal is singular if it is not regular.*

Definition A.2.7 (κ-like model)**.** *Let κ be a cardinal. We call a model M κ-like if, for all $a \in M$, we have that $|\{x \in M : x < a\}| < \kappa$ and $|M| = \kappa[12, Exercise 7.12]$.*

We now demonstrate that the cardinal ω is regular and that the model $\mathbb{N}$ is $\omega - like$.

Example A.2.8. *The cardinal $\omega = |\mathbb{N}|$ is regular and the standard model $\mathbb{N}$ is ω-like.*

Proof. Suppose $|X| = \omega$ and write X as $\bigcup_{i \in I} Y_i$. If both $|I| < \omega$ and $|Y_i| < \omega$ for every $i \in I$, then I is finite and every Y_i is finite. We know the finite union of finite sets is finite, and X is not finite. Thus $|I| = \omega$ or $Y_i = \omega$ for some $i \in I$.

Let $a \in \mathbb{N}$. We have $\{x \in \mathbb{N} : x < a\} = \{1, 2, ..., a-1\}$ and hence is finite, so has cardinality strictly smaller than ω. □

Further examples of regular cardinals are ω_n for each $n \in \omega$.

A.2.3 Recursive Saturation

The notion of a model being recursively saturated is an important one, which we define below.

Definition A.2.9. *A model M in the signature of some finite language $\mathscr{L}$ is recursively saturated if for any recursive set of (Gödel-numbers of) formulas*

$\{\ulcorner\theta_i(x,\bar{a})\urcorner : i \in \mathbb{N}\}$, *where* $\bar{a}$ *has finite length, it is the case that* $M \vDash \exists x \bigwedge_{i<n} \theta_i(x,\bar{a})$ *for each* $n \in \mathbb{N}$, *then* $M \vDash \exists x \bigwedge_{n\in\mathbb{N}} \theta(x,\bar{a})$.

A sentence of the form $\exists x \bigwedge_i \theta_i(x,\bar{a})$ *where* $n \in \mathbb{N}$ *is called a type [12, Definition 11.6].*

The following two theorems are key ones when looking at the theory of recursively saturated models.

Theorem A.2.10. *If* M *is a model in the signature of some finite language* $\mathscr{L}$, *then there exists a recursively saturated model* N *in the signature of* $\mathscr{L}$ *such than* $M \prec N$ *[12, Proposition 11.8].*

Theorem A.2.11. M *is not recursively saturated if and only if there is a recursive set of (Gödel-numbers of) formulas* $\{\ulcorner\theta_i(x,\bar{a})\urcorner : i \in \mathbb{N}\}$ *such that the set* $\{B_i : i \in \mathbb{N}\}$ *forms a partition of* M, *where* $B_i = \{x \in M : M \vDash \theta_i(x,\bar{a})\}$ *for each* $i \in \mathbb{N}$ *[12, Theorem 15.7].*

Appendix B

Literature Review

The aim of this thesis is to explore definitions of truth in models of arithmetic. This has been widely discussed within the literature, with many different starting approaches. A thorough analysis of all of these is beyond the scope of this work, and so we shall narrow our approach by starting with Tarski's inductive definition of truth for models of arithmetic [20].

Through a series of technical manoeuvres Tarski is able to define truth for atomic formulas in the language of arithmetic $\mathscr{L}_A$ [20, Section 5]. Tarski then introduces axioms which define how truth behaves with respect to connectives and quantifiers [20, Section 5]. Whilst this defines truth for all standard sentences of arithmetic, Tarski's approach is lacking in two areas. Firstly, the truth predicate can only be applied to sentences not containing the truth predicate and secondly, it does not tell us the truth value of sentences of nonstandard length. Tarski proves that this first weakness is a facet of all suitable definitions of truth for arithmetic, and is known as Tarski's Theorem on the Undefinability of Truth [20, Theorem 1].

Krajewski [14], using an idea from Robinson [17], takes up the second weakness and uses Tarski's axioms to define the arithmetic structures of satisfaction classes. These are definitions of truth for nonstandard models of arithmetic and he proves that if a countable model of arithmetic posesses a satisfaction class, then it has $2^{\aleph_0}$ satisfaction classes [14]. Lachlan then investigates satisfaction classes proves that a model of arithmetic can have a satisfaction class only if it is recursively saturated [15].

These theorems tell us that satisfaction classes carry ontological force as a mathematical structure and that satisfaction classes are not sufficient to provide one canonical definition of truth for models of arithmetic. As a corollary to this, there are satisfaction classes which state intuitively-true sentences are false, which we shall call pathological examples.

It could be argued that these consequences of satisfaction classes conflict with the philosophical conception of truth known as deflationism. The deflationist conception of truth been proposed by Frege [6], Ayer [2] and Quine [16], among others, but most clearly and recently by Horwich [8]. Horwich proposes that a sentence is true if and only if the meaning of the sentence holds and that truth has no metaphyiscal nature [8].

It is an interesting question as to whether these differing conceptions of truth can be reconciled and, if not, which one is more suited to models of arithmetic. Ketland [13] explores whether there is a conflict between the Tarskian definition of truth and the deflationary conception of truth with regards to arithmetic, and concludes that there is. Ketland's methods, however, are based on a stronger form of Tarskian truth [13, Pages 79-80] than proposed by Tarski. There is therefore a gap in the literature to explore this question more formally.

Ciéslińki [4] tackles the question of whether satisfaction classes can be reconciled with the deflationary conception of truth, but stresses that pathological examples must first be removed from satisfaction classes. He concludes that strengthening satisfaction classes to avoid pathological examples via closing them under propositional logic results in a theory which is equivalent to all induction axioms holding for Δ_0 formulas in the language of satisfaction classes, $\Delta_0 - \mathrm{PA}(S)$ [4, Theorem 4]. Therefore, Ciéslińki argues that proceeding in this way results in suitable satisfaction classes that are not deflationary and poses the open question of whether we can find similarly suitable satisfaction classes that align more closely with a deflationist conception of truth [4, Page 5]. This is a key avenue of research that this thesis shall endeavour to explore.

This question is explored by Kaye [12] who gives a presentation of satisfaction classes in languages including functions. Kaye provides examples of the pathologies which can exist within satisfaction classes [12, Page 251] and explores approaches to removing them. He shows that the addition of induction axioms to satisfaction classes is very powerful and that the theory $\Delta_0 - \mathrm{PA}(S) \vdash con(\mathrm{PA})$ [12, Page 247]. The literature is therefore lacking in methods which remove pathologous examples from satisfaction classes, without creating a very strong theory which conflicts with a deflationist conception of truth.

Avenues to dealing with this issue can be found in Kaye [12] and Visser and Enayat [5]. Kaye discusses the collection axiom schema and proves many interesting theorems about them. In particular, the collection axiom schema is strictly weaker than the induction axiom schema and is often very different to it [12, Chapter 7]. This opens up the approach of adding the collection axiom schema to the theory of satisfaction classes, in the hope that this produces a conservative theory that nevertheless removes pathologous examples. Kaye, further, observes that the collection axioms are linked to end-extensions of models [12, Section 7.2]. Visser and Enayat [5, Lemma 3.1] provide a new method of constructing satisfaction classes for models of arithmetic that produces an extension of a model and satisfaction class. It is therefore a suitable line of enquiry to see whether this approach can be improved to produce an end extension and, if so, whether this provides collection axioms for the satisfaction class.

This survey of the literature leads to the aim for this thesis to explore the current theory of collection axioms and satisfaction classes in detail, in order to facilitate an attempt at adding collection axioms to the theory of satisfaction classes. It is hoped that this will provide a definition of truth for models of arithmetic that is suitable in the sense that it does not contain pathologous examples, whilst not being as strong as $\Delta_0 - \mathrm{PA}(S)$, as per Ciéśliński's open question [4, Page 5]. The aim then proceeds to exploring the question of the compatibility of a deflationist conception of truth and satisfaction classes and the hope that progress can be made in light of this new background theory.

Appendix C

Deflationary Truth

C.1 Introduction

The deflationary theory of truth, which we shall often call deflationism, has many different accounts all with one common assertion. Deflationism argues that asserting a statement is true is equivalent to asserting the meaning of the statement itself. For example saying that the sentence ‘grass is green’ is true is equivalent to saying that grass is green. For the deflationist, there is nothing more to truth and to say that truth has a nature beyond this is incorrect. Throughout this section, we shall use single-quotation marks ‘ and ’ around a sentence to denote an expression of that sentence in natural language and hooked-quotation marks $\ulcorner$ and $\urcorner$ around a sentence to denote a numerical coding of a formal-language sentence. For now, this distinction is not an important one, but more on this can be seen in Subsection E.1.1.

C.2 Various Accounts of Deflationism

Deflationary theories of truth have been proposed in different guises by many philosophers. We shall explore a few of these accounts of deflationary truth below, picking out their similarities and also some of their differences, in order to attempt to formalise what makes a theory of truth deflationary.

C.2.1 Frege

Frege argued for a deflationist account of truth and it is in his writings we see the beginnings of the deflationist account of truth. He contends [6, Page 293] that "nothing is added to a thought by my ascribing to it the property of truth." For Frege, given a "thought" p, the thought 'p is true' contains no extra information. This tells us that there is no content to 'p is true' other than p.

Frege [6, Page 294] considers two counter-examples to this account of truth and contends that these are not actually counter-examples. The first he considers is that for the scientist the sentence 'theory T is true' adds to the original proposition of the theory T, where at proposal the theory's truth is in doubt. He argues that in these types of sentences, thinking sentences, a thought is apprehended without being asserted. A scientist's explanation of theory T is only apparent assertion, and in actual fact the thought has not been judged as true or false, so has not been proposed.

Frege [6, Page 295] also considers poetical sentences, which contain components over which truth does not extend. For example, the sentence 'alas my arm is bruised' conveys more information than the sentence: ' 'my arm is bruised' is true', however we cannot assert 'it is true that 'alas my arm is bruised.' ' There appears to be an information gap in the sentences. For Frege, this is not a counter-example to his deflationist theory of truth, since the word 'alas' does not actually assert anything, it instead acts on the feelings of the reader, and can not be expressed in any thought. Therefore, for Frege, there is nothing more to truth than the assertion of a thought, and the role it plays in our language is to highlight at which point a sentence becomes a thought.

C.2.2 Ayer

Another historical deflationist account can be found in the writings of A. J. Ayer. Arguing similar to Frege, Ayer writes [2, Page 28] that "it is evident that in a sentence of the form 'p is true' or 'it is true that p' the reference to truth never adds anything to the sense." This is the same account as found in Frege, where saying a thought is true containts nothing other than the sense of the thought.

Ayer adds to this deflationist account of truth by writing [2, Pages 28–29] that "*truth* and *falsehood* are not genuine concepts." Here, Ayer states that truth can be understood simply as assertion and does not have any other qualities, unlike genuine concepts which assign a quality to an object. For Ayer, there is no metaphysical nature to truth and it should not be considered as a genuine property. Theories of truth are not an account of a special concept, but instead are attempts to answer the empirical question as to how to validate propositions. Here we see in Ayer the two main constituents of a deflationist account of truth, that 'p' is true if and only if p and that truth has no special quality and this is all we understand by the concept of truth.

C.2.3 Quine

Quine also expressed a deflationist account of truth. He argues that "by calling the sentence ['snow is white'] true, we call snow white. The truth predicate is a device of disquotation" [16, Page 12]. Here Quine argues that given an expression of a proposition, e.g. 'snow is white', assigning truth to this allows us to remove the expression and just assert the proposition. This is the same deflationist account of truth that can be found in Frege and Ayer, that truth of sentences entails the truth of the meaning of the sentence.

One objection to the deflationist account of truth could be that if truth has no nature, then it has no utility in our language. Quine argues against this and writes that the reason for the truth predicate is in the affirmation of infinitely-many sentences. He gives the example that the truth predicate is required to affirm the statement "every sentence of the form 'p or not p' is true" [16, Page 12]. We have difficulty translating this without the word true. Thus, Quine argues, that if we were to express this as "p or not p for all things p of the sort that sentences are names of" [16, Page 11] then we are using p in two different ways, as a variable representing sentence clauses and representing substantive nouns. Hence, for Quine, the truth predicate's utility is in dealing with sentences such as this.

C.2.4 Horwich

Horwich is a contemporary deflationist who again argues for equivalence of a sentence being true if its meaning holds. He writes "for any declarative sentence (4) p we are provided with an equivalent sentence (4*) the proposition *that* p is true ... the truth predicate serves merely to restore the structure of a sentence: it acts simply as a de-nominalizer." [8, Pages 4-5]. This is a familiar account which we see is common to all deflationary theories of truth.

Howrich also argues against the misconception that "truth *has* some hidden structure awaiting our discovery" [8, Page 2]. He argues, like Ayer, that truth is not an ordinary property and writes, like Quine, the concept of truth only exists to fulfil a need to refer to an infinite conjunction of sentences. He argues that truth has no underlying nature and the deflationary principle that truth is a de-nominaliser is the only explanation required for truth. This principle captures everything we need to know about truth and provides everything we need from an account of truth.

These two principles contain the essence of the deflationist account of truth. They claim to provide a necessary condition for truth to satisfy, which is also a sufficient one. Whilst there is a lot more to say about the deflationist account of truth, this feature is specific enough for our interests.

C.2.5 Tarski

Tarski has provided a highly interesting conception of truth which provides the foundation for satisfaction classes. We shall examine this conception in Section E.1. It is an interesting question as to the extent to which Tarski was a deflationist about truth, and whether his account can be considered deflationary or not. We shall consider this in Subsection E.1.5 after a thorough account of his definition of truth.

C.3 Meta and In-Model Truth

We have seen various different deflationary accounts of truth, all of which have one very important facet in common. They all agree that, for example, ‘snow is white’ is true if and only if snow is white and this is all we understand by the nature of truth. More generally, for a sentence p, it is held that ‘p’ is true if and only if p. This definition of deflationary truth is known as the equivalence schema. We can express this in our natural language in the following way:

$$\langle p \rangle \text{ is true if and only if } p.$$

where p is a proposition and $\langle p \rangle$ is the representation of that proposition.

We see that the equivalence schema describes truth in a very general way which applies to everyday language. Given any reasonable sentence of English, this schema should give an account of truth for these sentences. These sentences reasonably include all mathematical sentences as well. Applying the deflationary account of truth to mathematical sentences we get an account of truth which says ‘$2+2=4$’ is true if and only if $2+2=4$. We hence have the equivalence schema that $\mathrm{Tr}^m(\delta)$, that ‘δ’ is true, if and only if δ for a sentence (of potentially natural language) δ. We shall call this account of truth a meta-truth account and denote it by the truth predicate Tr^m.

This is not the only way we can conceive of truth within mathematics. We can instead look at truth from the perspective of inside a model. Rather than generally specifying the truth of a sentence, we specify the truth of a sentence from the perspective of a mathematical model. This truth definition, particular to a specific model, will change based on the model we are considering. We shall call this account of truth a model-truth account and, for a model M we shall denote it by the truth predicate Tr^M. We can apply a deflationary account of truth to this model-truth notion as well. We thus have, for all models M that $M \vDash \mathrm{Tr}^M(\sigma)$ if and only if $M \vDash \sigma$ for all sentences σ. We therefore have two different equivalence schemas to consider.

It is an interesting question as to the dependencies between the meta-truth equivalence schema and the model-truth equivalence schema. The meta-truth account as proposed is not sufficient for the truth of mathematical sentences involving quantified variables. For instance, it is true in $\mathbb{N}$ that $\forall x[x \geqslant 0]$, but the same sentence is false in $\mathbb{Z}$. Thus, it is infeasible to write $Tr^m(\forall x[x \geqslant 0])$. This problem of not qualifying mathematical sentences is exacerbated further for the mathematical realist. For instance, we know that $\mathbb{N}$ is consistent, not just in the sense of a model but actually consistent, and therefore it could appear reasonable to write $\mathrm{Tr}^m(con(\mathbb{N}))$. We also have that $\mathbb{N} \vDash \mathrm{PA}$ and therefore write $\mathrm{Tr}^m(\mathbb{N} \vDash \mathrm{PA})$. Given the previous two statements one may be tempted to write $\mathrm{Tr}^m(con(\mathrm{PA}))$. However, from Gödel's second incompleteness theorem we know that there are models $M \vDash \mathrm{PA}$ such that $M \vDash \neg con(\mathrm{PA})$. Therefore, for some M, we have that $M \vDash \mathrm{Tr}^M(\neg con(\mathrm{PA}))$. This tells us that we cannot simply 'push' meta-truth inside a model for a model-truth and likewise cannot simply 'pull' a model-truth outside of a model to give us meta-truth and there is a disconnect between the two notions. Therefore, we shall consider the meta-truth predicate only to apply to sentences of the form $M \vDash \sigma$ where M is a model and σ is a sentence in the signature of the model.

It is therefore, very desirable, after specifying this is how meta-truth behaves, to want the equivalence that $\mathrm{Tr}^m(M \vDash \sigma)$ if and only if $M \vDash \mathrm{Tr}^M(\sigma)$. We shall thus define a model-truth account as being a good account of truth if it satisfies this property.

C.4 Deflationary Model-Truth

We now look more closely at a deflationary model-truth notion and the model-theoretic consequences that it proposes with Ketland's formal definition of deflationary truth. We shall do this here within the background language of arithmetic, $\mathscr{L}_A$ and the theory of first order Peano Arithmetic, PA, but more general approaches can be taken. Ketland [13] provides the following definition.

Definition C.1 (Ketland)**.** *The truth theory T with predicate Tr^M is deflationary (in the sense of Ketland) if for all models $(M, \mathrm{Tr}^M) \vDash \mathrm{PA} + T$:*

$$(M, \mathrm{Tr}^M) \vDash \mathrm{Tr}^M(\ulcorner \sigma \urcorner) \leftrightarrow \sigma$$

for all sentences σ in the signature of M. Further, we require that T is conservative over PA*. [13, Pages 71–72]*

This is a definition of what it means for a predicate, Tr^M to be deflationary, but does not provide us with an actual truth predicate. The natural approach to providing a deflationary truth predicate is by giving the equivalence schema as the list of axioms for the predicate. We follow Ketland [13, Section 2] and define the deflationary account of truth known as the *disquotational theory* (DT) in this way.

Definition C.2. *DT consists of the predicate* Tr^M *with the axiom schema: for all $\mathscr{L}_A$ sentences σ: $Tr^M(\ulcorner\sigma\urcorner) \leftrightarrow \sigma$. [13, Page 75]*

This theory DT is indeed a deflationary theory of truth, as the following theorem shows. Clearly it satisfies the first part of the definition, for that is how we defined DT but it is less obvious that this theory is conservative.

Theorem C.3. *DT is a deflationary theory of truth for* PA*, in the sense that every model of* PA *has an expansion which satisfies DT.*

Proof. Let $M \vDash \mathrm{PA}$, we build an expansion $(M, Tr^M) \vDash \mathrm{PA} + DT$. We add constants c_σ for each $\mathscr{L}_A$-sentence σ to M and then take a Gödel-Numbering so $\ulcorner\sigma\urcorner = c_\sigma$ for all $\mathscr{L}_A$-sentences σ. We then specify $Tr^M(\ulcorner\sigma\urcorner)$ if and only if $M \vDash \sigma$. Clearly $(M, \mathrm{Tr}^M) \vDash \mathrm{PA} + DT$. Since we took $M \vDash \mathrm{PA}$ arbitrarily, we thus have that any model has an expansion to DT. [13, Pages 76–77] □

This theorem therefore shows that it is possible to provide a formal definition of what it means for a truth predicate to be deflationary and further that under some definitions there are truth predicates that satisfy this.

Appendix D

Collection Axioms

In this chapter we give a thorough account of the collection axioms. We define what the first order collection schema is and specify which models satisfy them. We then look at their numerous links to the first order induction schema and prove various equivalences and conservativity results. We then look at the first order collection schema and its relation to end extensions of models. After this, we move onto the collection axioms over countably many conjuncts and disjuncts and look at its relation to recursive saturation. Finally, we look at the second order collection axiom, as a demonstration of the surprising weakness of collection.

D.1 First Order Collection

D.1.1 Collection Axioms

We begin by defining the first order collection axioms. The natural language interpretation of this for some formula θ is that if for all x lower than some bound, then there is some y such that $\theta(x, y, \bar{a})$ holds, then there is some bound on y. We define this formally below.

Definition D.1.1 (Collection Axioms). *The collection axiom,* Coll_θ*, for each formula* θ *of our language* $\mathscr{L}$ *is the sentence:*

$$\forall \bar{a} \forall t [\forall x{<}t \exists y \theta(x, y, \bar{a}) \rightarrow \exists z \forall x{<}t \exists y{<}z \theta(x, y, \bar{a})]$$

where $\bar{a}$ *represents a (possibly empty) tuple of variables [12, Definition 7.2].*

Definition D.1.2 (Collection Schema). *For a given language* $\mathscr{L}$ *we denote* $\{\mathrm{Coll}_\theta : \theta$ *is an* $\mathscr{L}$*-formula*$\}$ *as* Coll*. Thus, we say that a model* M *satisfies* Coll, $M \vDash \mathrm{Coll}$*, if for each formula* θ *of* $\mathscr{L}$, $M \vDash \mathrm{Coll}_\theta$.

We shall also consider collection restricted to certain formulae classes. We define these as one might expect, below:

Definition D.1.3 ($\mathrm{Coll}(\Sigma_n)$ and $\mathrm{Coll}(\Pi_n)$). *For our language* $\mathscr{L}$ *we denote* $\{\mathrm{Coll}_\theta : \theta$ *is an* $\mathscr{L}$*-formula in the set* $\Sigma_n\}$ *by* $\mathrm{Coll}(\Sigma_n)$*. Similarly, by* $\mathrm{Coll}(\Pi_n)$*, we denote* $\{\mathrm{Coll}_\theta : \theta$ *is an* $\mathscr{L}$*-formula in the set* $\Pi_n\}$

We now explore the collection axioms in more detail, by looking at which models satisfy these axioms.

D.1.2 Models of Collection

It is clear to see that the standard model $\mathbb{N}$ of PA satisfies Coll. This is because, for a given t, we have finitely-many x which have some y such that $\theta(x, y)$. Therefore, there are finitely-many y which do this, so we can bound these y (e.g. take z as the maximum of these y, $+1$). It is not obvious that nonstandard models of PA should satisfy Coll, however. This is because for nonstandard t we have infinitely-many x and so infinitely-many y which may therefore be unbounded. It turns out that due to induction they are bounded and we shall prove this below.

Theorem D.1.4. *If* $M \vDash \mathrm{PA}$*, then* $M \vDash \mathrm{Coll}$ *[12, Proposition 7.4].*

Proof. We prove this via induction on t. First let θ be an $\mathscr{L}$-formula and a requisite tuple of constants $\bar{a}$ be given.

First suppose $t = 0$ and $M \vDash \forall x{<}t \exists y \theta(x, y, \bar{a})$. This is vacuously true, since there are no $x < 0$. Therefore, it is similarly vacuously true that $M \vDash \exists z \forall x{<}t \exists y{<}z \theta(x, y, \bar{a})$.

Now suppose inductively that:

$$M \vDash \forall x{<}n \exists y \theta(x, y, \bar{a}) \rightarrow \exists z \forall x{<}n \exists y{<}z \theta(x, y, \bar{a}).$$

Consider if $M \vDash \forall x{<}(n{+}1) \exists y \theta(x, y, \bar{a})$. Then $M \vDash \forall x{<}n \exists y \theta(x, y, \bar{a})$ and $M \vDash \exists y \theta(n, y, \bar{a})$. Thus, by induction, $\exists z \forall x{<}n \exists y{<}z \theta(x, y, \bar{a})$, denote such a z by u. We now denote a y satisfying $\theta(n, y, \bar{a})$ by v. Then, take $w = max(u, v + 1)$ and we have that $M \vDash \forall x{<}(n{+}1) \exists y{<}w \theta(x, y, \bar{a})$. Thus, $M \vDash \exists z \forall x{<}(n{+}1) \exists y{<}z \theta(x, y, \bar{a})$. □

Whilst every model of PA is a model of collection, there are numerous models of PA^- which do not satisfy collection, as well as many that do. It is interesting to explore some of these to see why collection appears and why it fails. We now prove that every model of PA^- that is ω_1-like satisfies collection.

Theorem D.1.5. *If $M \vDash \mathrm{PA}^-$ and M is ω_1-like, then $M \vDash$ Coll [12, Exercise 7.12].*

Proof. We prove this by contradiction. Suppose that $M \vDash \forall x{<}t \exists y \theta(x,y)$ and that $M \vDash \neg\exists z \forall x{<}t \exists y{<}z \theta(x,y)$. Then we know that $\bigcup_{x<t}\{y : \theta(x,y)\}$ is cofinal in M. This tells us that $M = \bigcup_{x<t}\{m < y : \theta(x,y)\}$. Since M is ω_1-like, we know that this is a countable union of countable sets, so countable. This is a contradiction, so thus $M \vDash$ Coll. □

This argument generalises to other regular cardinals, which we prove when we consider the second order collection axiom. For now, we begin by demonstrating models of PA^- that do not satisfy collection. Our first example is that $\mathbb{Z}[X]^+$ does not satisfy collection.

Example D.1.6. $\mathbb{Z}[X]^+ \vDash \mathrm{PA}^- + \neg\mathrm{Coll}$ *[12, Exercise 7.12].*

Proof. To generate $\mathbb{Z}[X]^+$ we take the integers, $\mathbb{Z}$, and generate its ring of polynomials $\mathbb{Z}[X]$. We then define the order on it that $p > q$ if and only if $p - q > 0$, where a polynomial $p = p_n X^n + p_{n-1} X^{n-1} + ... + p_1 X + p_0 > 0$ in $\mathbb{Z}[X]$ if and only if $p_n > 0$ in $\mathbb{Z}$. Then, we take $\mathbb{Z}[X]^+ = \{p \in \mathbb{Z}[X] : p \geqslant 0\}$ and it is easy to check that this structure satisfies PA^-.

We now show that it doesn't satisfy collection. Consider the definable formula $\theta(x,y)$ which says that: "the largest non-zero degree of y is the coefficient of the largest non-zero degree of x." Then we know $\forall x{<}X^2 \exists y \theta(x,y)$. E.g. for $x = 9X + 7$ we can take $y = X^9$. However, there is no polynomial z such that $\forall x{<}X^2 \exists y{<}z \theta(x,y)$. To see this suppose that such a

$z = z_n X^n + z_{n-1} X^{n-1} + ... + z_1 X + z_0$ exists. Then consider $x = (n+1)X < X^2$. This has corresponding $y \geqslant X^{n+1}$ and we see that hence $y > z$ which is contradicts that $y < z$. □

We can generalise this example easily enough to show that any model $M \vDash \mathrm{PA}^-$ has an end-extension $M[X]^+ \vDash \mathrm{PA}^- + \neg\mathrm{Coll}$. Whilst we do not prove this here, it is easily done using the above proof.

Theorem D.1.7. *If $M \vDash \mathrm{PA}^-$, then $M[X]^+ \vDash \mathrm{PA}^- + \neg\mathrm{Coll}$.*

These examples show us that whilst the collection axiom schema is true in every model of PA, there are numerous models of PA^- which do not satisfy it. Given that Theorem D.1.4 required induction to be proven, it is unsurprising that this is the case. The question of the link between collection and induction is an interesting one to explore, which we focus on in the next subsection.

D.1.3 Collection and Induction

The intuition that the collection schema and induction are related, as expressed above, is correct. We shall characterise this below, by proving that PA is equivalent to $\mathrm{I}\Delta_0 + \mathrm{Coll}$. We shall then see how this entails that Coll requires some induction to prove even some of the simplest theorems of PA, and for any Π_2 sentence it is actually conservative over $\mathrm{I}\Delta_0$. It is this dual nature of collection, of being very powerful with induction, but very weak on its own, that we focus on.

Our first lemma shows that using Π_n induction and Σ_{n+2} collection, we are able to build Σ_{n+1} induction. It is this lemma that is at the heart of our main proof later on, that PA is equivalent to $I\Delta_0 + \mathrm{Coll}$.

Lemma D.1.8. $I\Pi_n + \mathrm{Coll}(\Sigma_{n+2}) \vdash I\Sigma_{n+1}$, *where* $n \geqslant 0$.

Proof. Let $M \models PA^-$ and suppose $M \models I\Pi_n + \mathrm{Coll}(\Sigma_{n+2})$. Take $\theta(x, \bar{a})$ to be a Σ_{n+1} formula, where $\bar{a}$ is a tuple of constants in M. Suppose further that $M \models \theta(0, \bar{a}) \wedge \forall x[\theta(x, \bar{a}) \rightarrow \theta(x{+}1, \bar{a})]$. Thus we aim to show that $M \models \forall x \theta(x, \bar{a})$.

Let $t \in M$ and, since $\theta(x, \bar{a})$ is Σ_{n+1}, let $\exists y \phi(x, y, \bar{a})$ be an equivalent formula, where $\phi(x, y, \bar{a})$ is a Π_n formula. By the axioms of PA^-, we know that:

$$M \models \forall x{<}t+1(\exists y[\phi(x, y, \bar{a}) \vee \forall z \neg\phi(x, z, \bar{a})]).$$

This formula is Σ_{n+2} since $\forall z \neg\phi(x, z, \bar{a})$ is a Π_{n+1} formula. Hence we can apply collection to this sentence to get that, for some constant c:

$$M \models \forall x{<}t+1(\exists y{<}c[\phi(x, y, \bar{a}) \vee \forall z \neg\phi(x, z, \bar{a})]).$$

Thus, we see that $M \models \forall x{<}t+1[\exists y \phi(x, y, \bar{a}) \leftrightarrow \exists y{<}c\phi(x, y, \bar{a})]$. Here, we see that $\exists y{<}c\phi(x, y, \bar{a})$ is equivalent to some Π_n formula $\psi(x, c, \bar{a})$. Therefore, by $I\Pi_n$ we have that:

$$M \models \psi(0, c, \bar{a}) \wedge \forall x{<}t[\psi(x, c, \bar{a}) \rightarrow \psi(x{+}1, c, \bar{a})].$$

Hence, by performing Π_n induction on the formula $(x{>}t \vee (x \leqslant t \wedge \psi(x, c, \bar{a})$ we get that $M \models \psi(t, c, \bar{a})$. Therefore, $M \models \exists y \phi(t, y, \bar{a})$. Thus, we get that $M \models \theta(t, \bar{a})$. Therefore, since t was chosen arbitrarily, we get $M \models \forall x \theta(x, \bar{a})$.

[12, Lemma 7.6] □

We now require the following lemma, which tells us that Σ_n induction implies Π_n induction, which provides the other building block of our main theorem in this section. It is actually the case that Σ_n and Π_n induction are equivalent [12, Lemma 7.7], but as the other direction is not needed for our purpose, the proof is not given here.

Lemma D.1.9. $\mathrm{I}\Sigma_n \vdash \mathrm{I}\Pi_n$.

Proof. Let $M \models \mathrm{PA}^- + \mathrm{I}\Sigma_n$ and let $\phi(x, \bar{a})$ be a Π_n formula, where $\bar{a}$ is a tuple of constants in M. Suppose ($\star$) that:

$$M \models \phi(0, \bar{a}) \wedge \forall x[\phi(x, \bar{a}) \rightarrow \phi(x{+}1, \bar{a})].$$

We assume for contradiction that there exists some constant $c \in M$ such that $M \models \neg\phi(c, \bar{a})$. We now take $\psi(x, c, \bar{a})$ as the formula:

$$(x{>}c) \vee \exists y[y{+}x{=}c \wedge \neg\phi(y, \bar{a})].$$

We see that since $\neg\phi(y, \bar{a})$ is Σ_n we have that $\psi(x, c, \bar{a})$ is also a Σ_n formula. Since $M \models \neg\phi(c, \bar{a})$ we know that $c{+}0{=}c \wedge \neg\phi(c, \bar{a})$. Therefore we conclude that $M \models \psi(0, c, \bar{a})$. Now suppose that $M \models \psi(x, c, \bar{a})$. Then $x > 0$, in which case $x{+}1 > 0$, or for some b, $[b + x = c$ and $\neg\phi(b, \bar{a})]$. In this case $(b-1)+(x+1) = c$ and by the contrapositive of ($\star$) we have that $\neg\phi(b{-}1, \bar{a})$. Therefore, we know that:

$$M \models \forall x[\psi(x, c, \bar{a}) \rightarrow \psi(x + 1, c, \bar{a})].$$

Now since $\psi(x, c, \bar{a})$ is Σ_n, we apply induction to get that $M \vDash \forall x \psi(x, c, \bar{a})$. Therefore, we have that $M \vDash \psi(c, c, \bar{a})$, which allows us to conclude that $M \vDash \exists y[y{+}c{=}c \wedge \neg\phi(y, \bar{a})]$. Clearly, such a y, must be 0 and therefore $M \vDash \neg\phi(0, \bar{a})$, which contradicts out earlier assumption. Therefore, no such c can exist and $M \vDash \forall x \phi(x, \bar{a})$. [12, Lemma 7.7] □

It is these two lemmas which allow us to prove one of our most surprising theorems in this chapter, that PA is equivalent to $\mathrm{I}\Delta_0 + \mathrm{Coll}$. One of the main pieces of interest in this theorem is that it shows that collection, with just a small amount of induction, can be very powerful.

Theorem D.1.10. *The axioms of* PA *are equivalent to* $\mathrm{I}\Delta_0 + \mathrm{Coll}$ *[12, Theorem 7.5].*

Proof. We have seen in Theorem D.1.4 that $\mathrm{PA} \vdash \mathrm{Coll}$. Clearly, by definition $\mathrm{PA} \vdash \mathrm{I}\Delta_0$, so $\mathrm{PA} \vdash \mathrm{I}\Delta_0 + \mathrm{Coll}$.
Now consider $\mathrm{I}\Delta_0 + \mathrm{Coll}$. By definition $\mathrm{I}\Delta_0 = \mathrm{I}\Pi_0$ and by Lemma D.1.8 we thus have $\mathrm{I}\Delta_0 + \mathrm{Coll} \vdash \mathrm{I}\Sigma_1$. Now given $\mathrm{I}\Sigma_n$ we have by Lemma D.1.9 $\mathrm{I}\Pi_n$. So, again applying Lemma D.1.8 $\mathrm{I}\Sigma_n + \mathrm{Coll} \vdash \mathrm{I}\Sigma_{n+1}$. Thus, by induction, $\mathrm{I}\Delta_0 + \mathrm{Coll} \vdash \mathrm{I}\Sigma_n$ for all n. Therefore, $\mathrm{I}\Delta_0 + \mathrm{Coll} \vdash \mathrm{PA}$. □

Whilst we have seen that PA^- with collection only needs a little induction to become very powerful, without this induction even some 'simple' theorems of PA cannot be proven. The following example shows that $\mathrm{PA}^-{+}\mathrm{Coll}$ cannot prove that every number x lies between two numbers, $2y$ and $2(y+1)$.

Example D.1.11. $\mathrm{PA}^- + \mathrm{Coll} \nvdash \forall x \exists y[2y \leqslant x \wedge x < 2(y+1)]$ *[10, Exercise 7.7]*

Proof. We consider the model $\mathbb{Z}[X]^+$ and first show that:

$$\mathbb{Z}[X]^+ \nvDash \forall x \exists y [2y \leqslant x \wedge x < 2(y+1)].$$

We denote $[2y \leqslant x \wedge x < 2(y+1)]$ by the formula $\theta(x, y)$.

To demonstrate the claim we show that $\mathbb{Z}[X]^+ \vDash \neg\theta(X, y)$. Suppose $y \in M$ such that $\mathbb{Z}[X]^+ \vDash \theta(X, y)$. We write y as:

$$y_n X^n + y_{n-1} X^{n-1} + ... + y_1 X + y_0$$

where $y_n > 0$ and $y_i \in \mathbb{Z}$ for all $i \in \{0, 1, ..., n\}$. If $n = 2$, then $y > X$, so $2y = y + y > x + x \geqslant x$, so $\theta(X, y)$ does not hold. If $n = 0$, then $2(y + 1) = 2(y_0 + 1) < X$, so $\theta(X, y)$ does not hold. Thus $n = 1$, so $y = y_1 X + y_0$ and $2y = 2y_1 X + 2y_0 > X$, since $y_1 > 0$. Thus $\theta(X, y)$ does not hold for any $y \in \mathbb{Z}[X]^+$.

We now build a chain of models $M_0, M_1, , , , , M_{\omega_1}$ where $M_0 = \mathbb{Z}[X]^+$, $M_{i+1} = M_i[X_{i+1}]^+$ and $M_\lambda = \bigcup[i < \lambda] M_i$ for each limit ordinal λ. We stop at the first uncountable model M_{ω_1} which is ω_1-like. Clearly $M_i \subseteq_e M_\lambda$ for each $i < \lambda$ where λ is a limit ordinal by definition. We also have that $M_i \subseteq_e M_i[X_{i+1}]^+$. Let $a \in M_i$ and let $y \in M_{i+1}$, so we have that:

$$y = y_n X_{i+1}^n + ... + y_1 X_{i+1} + y_0 < a.$$

Then we know that $y_n = ... = y_1 = 0$, so $y = y_0 \in M_i$. Thus we have that $\mathbb{Z}[X]^+ \subseteq M_{\omega_1}$. We now show that M_{ω_1} is ω_1-like. To see that $|M_{\omega_1}| = \omega_1$, note that each M_i contains at least i elements by definition. Let $a \in$

M_{ω_1} and consider $X_a = \{x \in M_{\omega_1} : x < a\}$. Then X_a is a (strict) end-extension of M_{ω_1}, so is countable, otherwise this contradicts that M_{ω_1} is the first uncountable model. Thus we have that $M_{\omega_1} \vDash \text{Coll}$ by Theorem D.1.5.

We know that the axioms of PA^- are all of the $\forall\exists$-form. These axioms are preserved upwards in end-extensions by an easy exercise, so thus we have that $M_{\omega_1} \vDash \text{PA}^-$. Similarly, as we know that $\mathbb{Z}[X]^+ \vDash \forall y[2y \leqslant X_0 \rightarrow X \geqslant 2(y+1)]$ we know that this is also preserved upwards and hence:

$$M_{\omega_1} \vDash \forall y[2y \leqslant X_0 \rightarrow X \geqslant 2(y+1)].$$

Therefore, there is a model of PA^- and collection which satisfies $\neg\theta(X_0, y)$ and our theorem is proven. $\square$

We are able to prove the sentence $\forall x \exists y[2y \leqslant x \wedge x < 2(y+1)]$ using only $\text{I}\Delta_0$ induction and no collection. This gives an example that although collection and induction are related, they are also useful for different tasks and should not be viewed as too closely related.

Example D.1.12. $\text{PA}^- + \text{I}\Delta_0 \vdash \forall x \exists y[2y \leqslant x \wedge x < 2(y+1)]$.

Proof. Let $M \vDash \text{PA}^- + \text{I}\Delta_0$ and let $t \in M$, where $t > 0$. By $\theta(x,t)$ we denote the formula:

$$\exists y{<}t[2y \leqslant x \wedge x < 2(y+1)] \vee x \geqslant 2t.$$

Clearly we have that $M \vDash \theta(0,t)$ since $0 < t$ and $2 \cdot 0 = 0 \leqslant 0$ and by the axioms of PA^- we have that $0 < 2(0+1) = 1$. Now suppose that $M \vDash \theta(x,t)$. Therefore $\exists y{<}t[2y \leqslant x \wedge x < 2(y+1)]$ or $x \geqslant 2t$. If there is some $c \in M$ such that $2c \leqslant x \wedge x < 2(c+1)$, then either $2c < x$ and $x < 2(c+1)$ or $2c = x$

and $x < 2(c+1)$. In the first case, we have that:

$$2(c+1) \leqslant x+1 \wedge x+1 < 2(c+2).$$

In the second case we have that:

$$2c \leqslant x+1 \wedge x+1 < 2(c+1).$$

Alternatively, if $x \geqslant 2t$, then clearly $x+1 \geqslant 2t$. Therefore $M \vDash \theta(x+1,t)$. Hence, we can use Δ_0 induction, to get that $M \vDash \forall x \theta(x,t)$. Since t was chosen arbitrarily, we thus have that $M \vDash \forall t \forall x \exists y{<}t[2y \leqslant x \wedge x < 2(y+1)] \vee x \geqslant 2t$. Therefore, we know:

$$M \vDash \forall t \forall x{<}2t \exists y{<}t[2y \leqslant x \wedge x < 2(y+1)] \vee x \geqslant 2t.$$

This hence gives us that $M \vDash \forall x \exists y[2y \leqslant x \wedge x < 2(y+1)]$. □

More surprisingly, for any Π_2 sentence, we have that Coll(Σ_1)+ IΔ_0 is conservative over IΔ_0. This shows that to prove a large amount of sentences of arithmetic collection over Σ_1 sentences is unnecessary. Thus, we know that in some circumstances the collection axioms are very weak.

Theorem D.1.13. *If* $\mathrm{PA}^- + \mathrm{Coll}(\Sigma_1) + \mathrm{I}\Delta_0 \vdash \sigma$ *where* σ *is a* Π_2 *sentence, then* $\mathrm{PA}^- + \mathrm{I}\Delta_0 \vdash \sigma$ *[3, Page 260].*

Proof. Without loss of generality we can write σ in the form $\forall x \exists y \theta(x,y,\bar{a})$ for some formula $\theta(x,y,\bar{a})$ where θ is Δ_0. Suppose we have a model $M \vDash \mathrm{I}\Delta_0$ where $M \nvDash \sigma$. Therefore, for some $a \in M$ we know that $M \vDash \forall y \neg \theta(a,y)$.

Without loss of generality suppose that M is countable (by the Löwenheim-Skolem theorems - see A.1.2), recursively saturated (by A.2.10) and $a > 2$ (since otherwise consider $a' = a+2$). We now consider the type $p(x) := \{x > a^n : n \in \mathbb{N}\}$. Since M is recursively saturated and $p(x)$ is finitely realised, we know that there is some $b > a^n$ for each $n \in \mathbb{N}$. Thus we consider the cut:

$$I = sup\{a^n : n \in \mathbb{N}\} = \{m \in M : m < a^n \text{ for some } n \in \mathbb{N}\}.$$

Clearly $I \subset_e M$ and is closed under $+$ and $\cdot$. Therefore $I \models \mathrm{PA}^-$ and $I \models \forall y \neg\theta(a, y)$ since θ is Δ_0 and $a \in I$ and Π_1 formulas are preserved downwards. We can express induction for a Δ_0 formula $\phi(x, \bar{a})$ with the sentence:

$$\forall \bar{a}, t[\phi(0, \bar{a}) \wedge \forall x{<}t(\phi(x, \bar{a}) \rightarrow \phi(x+1, \bar{a})) \rightarrow \forall x \leqslant b\phi(x, \bar{a})].$$

This sentence is Π_1, so similarly is preserved and thus $I \models \mathrm{I}\Delta_0$.

We now prove that $I \models \mathrm{Coll}(\Sigma_1)$. Let $t \in I$ such that $I \models \forall x{<}t \exists y \phi(x, y, \bar{a})$ where $\phi(x, y, \bar{a})$ is a Σ_1 formula. Therefore, without loss of generality, we can write it as $\exists z \psi(x, y, z, \bar{a})$ for some Δ_0 formula $\psi(x, y, z, \bar{a})$. Hence, we know that $I \models \forall x{<}t \exists y \exists z \psi(x, y, z, \bar{a})$. Therefore, by pairing the y and z together, we have that $\exists y \exists z \psi(x, y, z, \bar{a})$ is equivalent to $\exists y \zeta(x, y, \bar{a})$ where ζ is a Δ_0 formula. Therefore, it is the case that:

$$I \models \forall x{<}t \exists y \zeta(x, y, \bar{a}).$$

Thus $M \models \forall x{<}t \exists y \zeta(x, y, \bar{a})$. We know that for all constantsl $b \in M$ which are not in I, $M \vDash \forall x{<}t \exists y{<}b \zeta(x, y, \bar{a})$. Thus, by overspill, since $M \vDash \mathrm{I}\Delta_0$, we know for some $b \in I$ that $I \vDash \forall x{<}t \exists y{<}b \zeta(x, y, \bar{a})$ and hence we have equivalently that $I \vDash \exists z \forall x{<}t \exists y{<}z \phi(x, y, \bar{a})$. Therefore, $I \vDash \mathrm{Coll}(\Sigma_n)$. This tells us that, if $M \vDash \mathrm{PA}^- + \mathrm{I}\Delta_0 + \neg\sigma$, then there is some $\mathrm{I} \vDash \mathrm{PA}^- + \mathrm{I}\Delta_0 + \mathrm{Coll}(\Sigma_n) + \neg\sigma$. Hence, by contrapositive, we achieve the desired result. □

D.1.4 Collection and Largeness Properties

The collection axioms can also be proven to be equivalent to Keisler's Axiom 5 for largeness properties in PA^-. This equivalence is of great use in the next section in proving the existence of an elementary end extension for any model of PA^- that satisfies Coll.

First we develop the useful shorthand of $\mathrm{Q}x$ which, informally, denotes 'there are unboundedly many x'.

Notation D.1.14 ($\mathrm{Q}x$)**.** *For any $\mathscr{L}$-formula $\theta(x, \bar{a})$, where $\bar{a}$ is a tuple of constants, we write $\mathrm{Q}x\theta(x, \bar{a})$ to mean the sentence: $\forall y \exists x{>}y \theta(x, \bar{a})$.*

We first prove the simple lemma below, that shows this 'quantifier' is implied by $\forall$ and implies $\exists$.

Lemma D.1.15. *Let θ be an $\mathscr{L}$-formula. We have that for any model $M \vDash \mathrm{PA}^-$ it is the case that $\forall x\theta(x)$ implies $\mathrm{Q}x\theta(x)$ and that $\mathrm{Q}x\theta(x)$ implies $\exists x\theta(x)$.*

Proof. Suppose a model $M \vDash \forall x\theta(x)$ and take $m \in M$ arbitrary. Then $M \vDash \exists x{>}m\theta(x)$. Since this holds for all $m \in M$ we know that $M \vDash \forall m \exists x{>}m\theta(x)$. Now suppose $M \vDash \mathrm{Q}x\theta(x)$. Then $M \vDash \forall y \exists x{>}y\theta(x)$ and in particular $M \vDash \exists x\theta(x)$. □

The shorthand Qx is an example of a largeness property. The following definition is an axiom of all largeness properties, as detailed in *Models by Games* [7]. In this thesis we shall just be looking at the largeness property Qx, so Keisler's Axiom 5 has been stated in terms of this.

Definition D.1.16 (Keisler's Axiom 5)**.** *For any $\mathscr{L}$-formula $\theta(x, y, \bar{a})$:*

$$\mathrm{Q}x\exists y\theta(x, y, \bar{a}) \rightarrow \exists y\mathrm{Q}x\theta(x, y, \bar{a}) \vee \mathrm{Q}y\exists x\theta(x, y, \bar{a}).$$

Here $\bar{a}$ is a tuple of constants [7, Page 175].

Since all models that we shall be considering are models of PA^- we are able to simplify this axiom, as the following lemma shows.

Lemma D.1.17. *In* PA^- *for an $\mathscr{L}$-formula $\theta(x, y, \bar{a})$, Keisler's Axiom 5 is equivalent to the sentence:* $\mathrm{Q}x\exists y{<}b\theta(x, y, \bar{a}) \rightarrow \exists y{<}b\mathrm{Q}x\theta(x, y, \bar{a})$. *We denote this sentence by S_θ [7, Lemma 6.1.6].*

Proof. Suppose $M \vDash \mathrm{PA}^-$ and $M \vDash$ Keisler's Axiom 5. Let $\theta(x, y, \bar{a})$ be an $\mathscr{L}$-formula, where $\bar{a} \in M$ and $M \vDash \mathrm{Q}x\exists y{<}b\theta(x, y, \bar{a})$. Then $M \vDash \mathrm{Q}x\exists y[\theta(x, y, \bar{a}) \wedge y{<}b]$. Therefore, by Keisler's Axiom 5, we know that:

$$M \vDash \exists y\mathrm{Q}x[\theta(x, y, \bar{a}) \wedge y{<}b] \vee \mathrm{Q}y\exists x[\theta(x, y, \bar{a}) \wedge y{<}b].$$

The second disjunct is impossible to obtain, since if this occurred it would imply that $M \vDash \forall z\exists y{>}z\exists x[\theta(x, y, \bar{a}) \wedge y{<}b]$, which entails that:

$$M \vDash \exists y{>}b\exists x[\theta(x, y, \bar{a}) \wedge y{<}b].$$

This is a contradiction. Thus $M \vDash \exists y \mathrm{Q}x[\theta(x,y,\bar{a}) \wedge y{<}b]$, and so we have that:

$$M \vDash \exists y{<}b \mathrm{Q}x\theta(x,y,\bar{a}).$$

We now prove the converse direction. Suppose $M \vDash S_\theta$ for all $\mathscr{L}$-formulae $\theta(x,y,\bar{a})$ and $M \vDash \mathrm{Q}x\exists y\theta(x,y,\bar{a})$. If there is some $b \in M$ such that $M \vDash \mathrm{Q}x\exists y{<}b\theta(x,y,\bar{a})$, then by S_θ we have that $M \vDash \exists y{<}b\mathrm{Q}x\theta(x,y,\bar{a})$ and thus $M \vDash \exists y\mathrm{Q}x\theta(x,y,\bar{a})$. Hence suppose this does not occur, and $M \vDash \neg\exists w\mathrm{Q}x\exists y{<}w\theta(x,y,\bar{a})$. Therefore we have that:

$$M \vDash \forall w\exists z\forall x{>}z\forall y{<}w\neg\theta(x,y,\bar{a}).$$

We take such a $z \in M$ and call it the constant c, then by rearranging, derive that $M \vDash \forall w\forall y{<}w\forall x{>}c\neg\theta(x,y,\bar{a})$. However, by substituting in c and rearranging our original supposition, we also know that $\exists y\exists x{>}c\theta(x,y,\bar{a})$. Therefore we know that $M \vDash \forall w\exists y \geqslant w\exists x{>}c\theta(x,y,\bar{a})$ and thus we derive that $M \vDash \mathrm{Q}y\exists x\theta(x,y,\bar{a})$. □

This leads us to one of the main results in this section. We show that, for any $\mathscr{L}$-formula θ, the sentences S_θ and Coll_θ are equivalent. To obtain this, we first prove the stronger statement that for any $n \in \mathbb{N}$, a model satisfies S_θ for all Σ_n sentences if and only if $\mathrm{Coll}(\Pi_n)$ is also satisfied in that model.

Lemma D.1.18. *In* PA^- *the sentence* S_θ *for all* Σ_n *formulae is equivalent to* $\mathrm{Coll}(\Pi_n)$ *[12, Exercise 7.3].*

Proof. First suppose that a model $M \vDash \mathrm{PA}^- + \mathrm{Coll}(\Pi_n)$. Let $\theta(x,y,\bar{a})$ be a Σ_n formula, where $\bar{a}$ is a tuple of constants in M. We suppose that

$M \vDash \mathrm{Q}x\exists y{<}t\theta(x,y,\bar{a})$ and for contradiction also that $M \nvDash \exists y{<}t\mathrm{Q}x\theta(x,y,\bar{a})$. Then $M \vDash \forall y{<}t\exists z\forall x{>}z\neg\theta(x,y,\bar{a})$. Since $\theta(x,y,\bar{a})$ is a Σ_n formula we know that $\forall x{>}z\neg\theta(x,y,\bar{a})$ is equivalent to a Π_n formula. Thus, we can apply $\mathrm{Coll}(\Pi_n)$ to the above sentence and get that:

$$M \vDash \exists w\forall y{<}t\exists z{<}w\forall x{>}z\neg\theta(x,y,\bar{a}).$$

Hence there is some $c \in M$ such that $M \vDash \forall y{<}t\forall x{>}c\neg\theta(x,y,\bar{a})$. However, we also know that $M \vDash \mathrm{Q}x\exists y{<}t\theta(x,y,\bar{a})$, which entails that $M \vDash \exists x{>}c\exists y{<}t\theta(x,y,\bar{a})$. This is a contradiction, so therefore $M \vDash S_\theta$ for all Σ_n formulae.

We prove the converse by contrapositive. Suppose a model $M \vDash \mathrm{PA}^- + S_\theta$ for all Σ_n formulas θ. Consider a Π_n formula $\theta(x,y,\bar{a})$ where $\bar{a}$ is a tuple of constants in M and suppose $M \vDash \neg\exists z\forall x{<}t\exists y{<}z\theta(x,y,\bar{a})$. Therefore, $M \vDash \forall z\exists x{<}t\forall y{<}z\neg\theta(x,y)$. Applying Lemma D.1.15 we thus have that:

$$M \vDash \mathrm{Q}xz\exists x{<}t\forall y{<}z\neg\theta(x,y).$$

We can apply $S_{\forall y<z\neg\theta(x,y)}$, since this is equivalent to a Σ_n formula, to get that $M \vDash \exists x{<}t\mathrm{Q}xz\forall y{<}z\neg\theta(x,y)$. By using the properties of negation this tells us that $M \vDash \neg\forall x{<}b\exists w\forall z{>}w\exists y{<}z\theta(x,y)$. Thus, we conclude that $M \vDash \neg\forall x{<}t\exists y\theta(x,y)$. Hence, we have that $M \vDash \mathrm{Coll}(\Pi_n)$. $\square$

This allows us to show that any model of PA^- considers S_θ and Coll_θ equivalent for any $\mathscr{L}$-formula θ.

Theorem D.1.19. *For* PA^- *we have that* Coll *is equivalent to* S_θ*, for all* *formulae* θ.

Proof. First suppose that $M \vDash \text{Coll}$. We know that any $\mathscr{L}$-formula θ is Σ_n for some n. We know that $\text{Coll}(\Pi_n)$ holds, so therefore by Lemma D.1.19, S_{Σ_n} holds and hence S_θ holds for any formula θ. Now suppose that $M \vDash S_\theta$ for all $\mathscr{L}$-formula θ. A given $\mathscr{L}$-formula θ is Π_n for some n. We know that S_θ holds for all Σ_n formulas, so by Lemma D.1.19 we have that $\text{Coll}(\Pi_n)$ holds, so Coll_θ thus holds also. This is true for any formula, so $M \vDash \text{Coll}$. $\square$

D.1.5 Collection and End Extensions

One of the many uses of the collection axioms is in their relation to end extensions. Informally one can see how collection can be linked to elementary end extensions. Supposing that some sentence of a model M does not satisfy collection, i.e. the set of y considered is unbounded, then by building an end extension onto this model, we see that some new element shall act as a bound for these y and thus this sentence shall satisfy collection in the new model. We shall first prove that only a little bit of elementarity in and end extension guarantees some collection. Then, we shall prove that if our model of PA^- has a full elementary end extension, then it satisfies full collection. Finally, we have that surprisingly, a converse to this nearly exists, and if a countable model of PA^- satisfies collection, then it has an elementary end extension.

Lemma D.1.20. *If* $M, N \vDash \text{PA}^-$ *and* $M \prec_{e,\Sigma_n} N$, *then* $M \vDash \text{Coll}(\Sigma_n)$, *for each* $n > 0$ *[12, Exercise 7.14].*

Proof. Suppose $M \prec_{e,\Sigma_n} N$ and further suppose that $M \models \forall x{<}t \exists y \theta(x, y)$ where $\theta(x, y)$ is a Σ_n formula. Thus, suppose without loss of generality that $\theta(x, y)$ is the formula $\exists w \phi(x, y, w)$ where $\phi(x, y, w)$ is Π_{n-1}. Hence $M \models \forall x{<}t \exists y \exists w \phi(x, y, w)$. We can rewrite this equivalently as:

$$M \models \forall x{<}t \exists \langle y, w\rangle \psi(x, \langle y, w\rangle)$$

where $\psi(x, \langle y, w\rangle)$ is some Π_{n-1} formula that holds if and only if $\phi(x, y, w)$ holds. Thus, by Σ_n elementarity we have, since $t \in M$, that:

$$N \models \forall x{<}t \exists \langle y, w\rangle \psi(x, \langle y, w\rangle).$$

Therefore, by taking a $z \in N \backslash M$ we have that:

$$N \models \exists z \forall x{<}t \exists \langle y, w\rangle{<}z \psi(x, \langle y, w\rangle).$$

Thus, again by Σ_n elementarity, we have that:

$$M \models \exists z \forall x{<}t \exists \langle y, w\rangle{<}z \psi(x, \langle y, w\rangle).$$

Now, we can repeat our first steps in reverse, to rewrite this as:

$$M \models \exists z \forall x{<}t \exists y{<}z \theta(x, y).$$

□

The following theorem is thus easily obtained by the above lemma and is highly interesting in its own right.

Theorem D.1.21. *If a model $M \models \mathrm{PA}^-$ and there is some model N such that $M \prec_e N$, then $M, N \models \mathrm{Coll}$.*

Proof. Every sentence of our language $\mathscr{L}$ is Σ_n for some n. If $M \prec_e N$, then $M \prec_{e,\Sigma_n} N$ for each n. Thus, by Lemma D.1.20 $M \models \mathrm{Coll}(\Sigma_n)$ for each n. Hence $M \models \mathrm{Coll}$. Therefore, since $M \prec_e N$, we also have that $N \models \mathrm{Coll}$. $\square$

Surprisingly, a converse to the result above exists for countable models, which is a useful alternative to the MacDowell-Specker theorem, Theorem A.1.1. This result requires the theorem we proved in the previous section, Theorem D.1.19.

Theorem D.1.22. *If a countable model $M \models \mathrm{PA}^-$ and $M \models \mathrm{Coll}$, then there exists a model N such that $M \prec_e N$ [10, Exercise 8.10].*

Proof. Let $\mathscr{L}$ be the language of our model and take $\mathscr{L}^+ = \mathscr{L} \cup M \cup \{\infty\}$ which is $\mathscr{L}$ with a constant symbol c_a for each $a \in M$ and a new constant ∞. Let $T_0 = Th(M, a)_{a \in M}$ which is the elementary diagram of M, the set of all true sentences of M using the constants from $\mathscr{L} \cup M$. We build new theories T_n successively and then take our final theory $T = \bigcup_{n \in \mathbb{N}} T_n$. We denote by τ_n the conjunction of all sentences in T_n. We build our new theories inductively with the induction hypothesis: $H(n) = M \models \mathrm{Q}x\tau_n(x, \bar{a})$ where $\mathrm{Q}x$ is in the sense of Notation D.1.14 above. We now enumerate all $\mathscr{L}^+$ sentences $\sigma(\infty, \bar{a})$ and all pairs of $\mathscr{L}^+$ formulae $\phi(\infty, x)$ and elements $a \in M$ which can be done since M is countable.

Given a theory T_n and a sentence $\sigma(\infty, \bar{a})$ we want to add either σ or $\neg\sigma$ to T_n to create T_{n+1}. We now show that this is possible and that:

$$M \vDash \mathrm{Q}x[\theta(x, \bar{a}) \wedge \tau_n(x, \bar{a})] \text{ or } M \vDash \mathrm{Q}x[\neg\theta(x, \bar{a}) \wedge \tau_n(x, \bar{a})].$$

Suppose $M \nvDash \mathrm{Q}x[\theta(x, \bar{a}) \wedge \tau_n(x, \bar{a})$ and $M \nvDash \mathrm{Q}x[\neg\theta(x, \bar{a}) \wedge \tau_n(x, \bar{a})$. Then since for all $x \in M$ either $\theta(x, \bar{a})$ or $\neg\theta(x, \bar{a})$ holds, we know that if $M \nvDash \mathrm{Q}x\theta(x, \bar{a})$ then $M \vDash \exists y \forall x{>}y \neg\theta(x, \bar{a})$, so $M \vDash \mathrm{Q}x\neg\theta(x, \bar{a})$. Thus $M \nvDash \mathrm{Q}x\tau_n(x, \bar{a})$ which contradicts our induction hypothesis.

Given a theory T_n and a pair $\phi(\infty, x)$ and $a \in M$ we want to add $\forall y{<}a \neg\phi(\infty, y)$ or, if we cannot, we add $\phi(\infty, b) \wedge b{<}a$, where b is a constant in M. To show this is possible, suppose $M \nvDash \mathrm{Q}x[\forall y{<}a \neg\phi(x, y) \wedge \tau_n(x)]$. Thus, we know by our induction hypothesis that:

$$M \vDash \mathrm{Q}x \exists y{<}a[\phi(x, y) \wedge \tau_N(x)].$$

Now, by Lemma D.1.19 we have that $M \vDash \exists y{<}a \mathrm{Q}x[\phi(x, y) \wedge \tau_n(x)]$. Thus, taking such a $b \in M$ we get that $M \vDash \mathrm{Q}x[\phi(x, b) \wedge b{<}a \wedge \tau_n(x)]$. Thus, we can consistently add $\phi(\infty, b) \wedge b{<}a$ for some constant $b \in M$.

We dovetail the above two processes together to construct our theory T which is both consistent and complete. If T were inconsistent, then $T \vdash \bot$ and thus $T_n \vdash \bot$ for some $n \in \mathbb{N}$. Therefore we know that $M \vDash \mathrm{Q}x\bot$ which is false, since M is contradiction-free. We also have ensured that $\infty \neq m$ for any $m \in M$, since otherwise we have $M \vDash \mathrm{Q}x(\infty = x)$ which is clearly false.

We now, consider the types:

$$p_a(x) = \{x < a\} \cup \{x \neq b : b < a\}.$$

Suppose that $T + \exists x{<}a\phi(\infty, x)$ is consistent. Then by the construction above we know that $\forall x{<}a\neg\phi(\infty, x)$ was not added to T. Thus there is some $b \in M$ such that $\phi(\infty, b) \wedge b < a$ was added. Thus $T \vdash \exists x{<}a[\phi(\infty, x) \wedge x = b]$ and hence we know that:

$$T \nvdash \forall x{<}a[\phi(\infty, x) \rightarrow x \neq b].$$

Therefore we know the type $p_a(x)$ is not isolated. We can now apply the omitting types theorem, Theorem A.1.7, that there is some model $N \models T$ where each $p_a(x)$ is omitted.

Thus we have that $N \succ_e M$. This is because $N \models Th(M, a)_{a \in M}$, N contains a new point $\infty \notin M$ and N omits the types $p_a(x)$ so contains no element not in M that is smaller than some element of M. □

D.2 Collection and Recursive Saturation

D.2.1 $\mathscr{L}^{rec}_{\omega_1\omega}$

To explore the link between the collection axioms and recursive saturation we work in an infinite language called $\mathscr{L}^{rec}_{\omega_1\omega}$. The useful property of this language is that we are allowed ω-many conjuncts and disjuncts in a sentence of $\mathscr{L}^{rec}_{\omega_1\omega}$.

Definition D.2.1. *We define the language $\mathscr{L}^{rec}_{\omega_1\omega}$ as an extension of $\mathscr{L}_A$ in* *which sentences may contain ω-fold conjunctions and ω-fold disjunctions. These sentences may only contain a finite number of quantifiers. Further, we require that all formulas of $\mathscr{L}^{\mathrm{rec}}_{\omega_1\omega}$ are recursive and that only finitely many variables may appear in a sentence of $\mathscr{L}^{\mathrm{rec}}_{\omega_1\omega}$ [9, Section 1.1]*

D.2.2 Recursive Saturation and other Properties

We now give the definition of recursive saturation in $\mathscr{L}^{rec}_{\omega_1\omega}$ below. For more details on recursive saturation in general, see Section A.2.3.

Definition D.2.2. *We call a model M recursively saturated in $\mathscr{L}^{rec}_{\omega_1\omega}$ if for any tuple of constants $\bar{a} \in M$ any recursive family of $\mathscr{L}_A$ formulas $\{\theta_i : i \in \mathbb{N}\}$ we have that:*

$$M \vDash \bigwedge_{n\in\mathbb{N}} \exists x \bigwedge_{i<n} \theta_i(x, \bar{a}) \to \exists x \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a}).$$

We now define two further properties of models of PA, in $\mathscr{L}^{rec}_{\omega_1\omega}$, below, both of which are consequences of recursive saturation. The first property we look at is short recursive saturation.

Definition D.2.3 (Short recursive saturation)**.** *Suppose that M is a model for a signature with $<$ and we have that $<$ is a linear order on M with no greatest element. Further, suppose that θ_n is a first order formula for each $n \in \mathbb{N}$. We say that M is short recursively saturated in $\mathscr{L}^{rec}_{\omega_1\omega}$ if for any tuple of constants $\bar{a} \in M$ and any constant $t \in M$ we have that:*

$$M \vDash \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a}) \to \exists x{<}t \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a}) \ [12, Exercise 11.6].$$

It is easy to see that if a model is recursively saturated, then it is short recursively saturated, but the converse does not hold in general. Short recursive saturation has the following interesting property, that does not hold for full recursive saturation, that it is preserved through elementary end extensions.

Lemma D.2.4. *Suppose that $M \prec_e N$. We have that M is short recursively saturated if and only if N is short recursively saturated [12, Exercise 11.6].*

Proof. First suppose that N is short recursively saturated and that:

$$M \models \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a}).$$

Then by elementarity $N \models \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a})$. Hence using N's short recursive saturation we have that $N \models \exists x{<}t \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a})$. Therefore, by elementarity, $M \models \exists x{<}t \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a})$, so M is also short recursively saturated.

Now suppose that M is short recursively saturated. We prove this direction via contrapositive, and suppose $N \models \neg\exists x{<}t \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a})$. In the first case $t \in M$, so we know that $M \models \neg\exists x{<}t \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a})$ by elementarity, and so by the contrapositive of short recursive saturation: $M \models \neg \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a})$. Therefore, again by elementarity,

$$N \models \neg \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a}).$$

In the second case $t \in N \backslash M$. Then we know by elementarity that $M \models \forall y \neg\exists x{<}y \bigwedge_{n\in\mathbb{N}} \theta_i(x, \bar{a})$. Therefore, by short recursive saturation in M we have that $M \vDash \forall y \neg \bigwedge_{n\in\mathbb{N}} \exists x{<}y \bigwedge_{i<n} \theta_i(x,\bar{a})$. Thus by elementarity:

$$N \vDash \forall y \neg \bigwedge_{n\in\mathbb{N}} \exists x{<}y \bigwedge_{i<n} \theta_i(x,\bar{a}).$$

This tells us in particular $N \vDash \neg \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x,\bar{a})$. Therefore N is also short recursively saturated. □

We now look at our second property, that of a model being tall.

Definition D.2.5 (Tall)**.** *Similarly to the above, suppose that M is a model for a signature with $<$ and we have that $<$ is a linear order on M with no greatest element. Further, suppose that θ_n is a first order formula for each $n \in \mathbb{N}$. We call M tall in $\mathscr{L}^{\mathrm{rec}}_{\omega_1\omega}$ if for any tuple of constants $\bar{a} \in M$ we have that:*

$$M \vDash \bigwedge_{n\in\mathbb{N}} \exists x \bigwedge_{\ulcorner\theta\urcorner<n} [\exists y\theta(y,\bar{a}) \to \exists y{<}x\theta(y,\bar{a})] \to \exists x \bigwedge_{\ulcorner\theta\urcorner} [\exists y\theta(y,\bar{a}) \to \exists y{<}x\theta(y,\bar{a})]$$

[12, Exercise 11.6].

For all models that we shall be considering the antecedent is always true. Thus, we can reduce the criterion to a model M being tall to:

$$\exists x \bigwedge_{\ulcorner\theta\urcorner} [\exists y\theta(y,\bar{a}) \to \exists y{<}x\theta(y,\bar{a})].$$

It is this notion of tall that we shall use throughout the rest of this section.

It is again easy to see that a recursively saturated model, satisfying the required properties for tallness, is tall. For any $n \in \mathbb{N}$ we know that if there exists an x such that for all formulae of Gödel-number less than n if $\exists y\theta(y)$, then $\exists y{<}x\theta(y)$. Hence, by recursive saturation, there is some x such that this holds for all formulae.

Interestingly, for models in which all Skolem-terms are definable (such as a model of PA) we can rephrase this in terms of Skolem functions. M is tall if $\exists x$ such that, for all Skolem-terms t, $t(\bar{a}){<}x$.

It is unsurprising that the natural numbers are not tall, we prove this below.

Example D.2.6. $\mathbb{N}$ *is not tall.*

Proof. Suppose that $\mathbb{N}$ is tall, ie:

$$\mathbb{N} \models \exists x \bigwedge_{\ulcorner\theta\urcorner} [\exists y \theta(y, \bar{a}) \to \exists y{<}x\theta(y, \bar{a})].$$

Then we know that such an x is a natural number, so is uniquely defined by a standard formula $\phi(x)$. Therefore $\exists y\phi(y)$ and by tallness we know that $\exists y{<}x\phi(y)$. Yet the only such possible y is x, telling us that $x < x$, which is a contradiction. So $\mathbb{N}$ cannot be tall. □

We have seen that both of these properties are consequences of recurisve saturation, but we also have the following useful fact which gives us the converse to this:

Fact D.2.7. *For any model $M \models$ PA we have that M is tall and short recursively saturated if and only if M is recursively saturated [12, Exercise 11.6].*

D.2.3 Collection, Short Recursive Saturation and Tallness

To start, we need to provide definitions of the collection axioms which utilises ω-many conjuncts or disjuncts. We then prove new results on dependencies between these definitions and tall and short recursively saturated models.

Definition D.2.8. *We define* Coll $\bigwedge$*, collection over ω-many conjuncts, as the set of all statements:*

$$\forall x{<}t\exists y \bigwedge_{n\in\mathbb{N}} \theta_n(x,y,\bar{a}) \rightarrow \exists z\forall x{<}t\exists y{<}z \bigwedge_{n\in\mathbb{N}} \theta_n(x,y,\bar{a})$$

for all tuples of constants $\bar{a} \in M$, all $t \in M$ and all first order formulae θ_n which are recursive.

Definition D.2.9. *We define* Coll $\bigvee$*, collection over ω-many disjunctions similarly. This is the set of all statements:*

$$\forall x{<}t\exists y \bigvee_{n\in\mathbb{N}} \theta_n(x,y,\bar{a}) \rightarrow \exists z\forall x{<}t\exists y{<}z \bigvee_{n\in\mathbb{N}} \theta_n(x,y,\bar{a})$$

for all tuples of constants $\bar{a} \in M$, all $t \in M$ and all first order formulae θ_n which are recursive.

There are numerous interesting interdependencies between Coll $\bigvee$, short recursive saturation and tallness, in particular a model satisfies Coll $\bigvee$ if and only if it is short recursively saturated or tall. Whilst in this section we focus on Coll $\bigvee$, for this is the property that was of use in Chapter 2, we also have a result concerning Coll $\bigwedge$ and tallness. To start with we repeat Lemma 2.1.7, that every model which satisfies Coll and is short recursively saturated also satisfies Coll$\bigvee$.

Lemma D.2.10. *If a model* $M \vDash$ Coll *and* M *is short recursively saturated, then* $M \vDash$ Coll$\bigvee$.

Proof. We prove this by contrapositive. Suppose M is a short recursively saturated model and that $M \vDash \neg\exists z \forall x{<}t \exists y{<}z \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$. Thus, we can rewrite this as:

$$M \vDash \forall z \exists x{<}t \forall y{<}z \bigwedge_{n\in\mathbb{N}} \neg\theta_n(x, y, \bar{a}).$$

This implies that $M \vDash \forall z \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \forall y{<}z \neg\theta_i(x, y, \bar{a})$. We can swap the universal quantification and conjunction at the beginning of this sentence, to get that $M \vDash \bigwedge_{n\in\mathbb{N}} \forall z \exists x{<}t \bigwedge_{i<n} \forall y{<}z \neg\theta_i(x, y, \bar{a})$. Now, by the contrapositive of Coll, we have that:

$$M \vDash \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \forall y \neg\theta_i(x, y, \bar{a}).$$

Now, using the fact that M is short recursively saturated we get that $M \vDash \exists x{<}t \bigwedge_{n\in\mathbb{N}} \forall y \neg\theta_n(x, y, \bar{a})$. Thus, by swapping the universal quantifier and the conjunction and by negation, we get that $M \vDash \forall x{<}t \exists y \bigwedge_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$. Therefore, by contrapositive, $M \vDash$ Coll$\bigvee$. □

The converse to this result does not hold, as the following example shows.

Example D.2.11. *There is a model* $M \vDash$ Coll$\bigvee$ *which is not short recursively saturated.*

Proof. Let $M \models \mathrm{PA}$ such that M is not short recursively saturated. For an example of this, we can consider $M = \mathrm{cl}_N(\varnothing)$, where $N \models \neg con(\mathrm{PA})$. For such a model we can let t be a nonstandard element, then we know that:

$$\mathrm{cl}_N(\varnothing) \models \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{\ulcorner\theta\urcorner<n} [\exists y\theta(y) \wedge \forall u\forall v(\theta(u) \wedge \theta(v) \rightarrow u{=}v) \rightarrow \neg\theta(x)].$$

What this says is that there is some standard number which is not defined by a standard formula of Gödel Number below any natural number. However, we do know that $\mathrm{cl}_N(\varnothing)$ satisfies the negation of this sentence, since all elements of $\mathrm{cl}_N(\varnothing)$ are definable.

Given a suitable model M, such as the one above, we take $M_0 = M$ and for every $i \geqslant 0$ given M_i we build an elementary end extension M_{i+1} such that if $M_i \models \forall x{<}t\exists y \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$, where t and $\bar{a} \in M_i$, then $M_{i+1} \models \exists z\forall x{<}t\exists y{<}z \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$. We know such end extensions exist from Theorem D.1.22. Then by taking $M_\omega = \bigcup_{i\in\mathbb{N}} M_i$ we have that $M_\omega \models \mathrm{Coll}\bigvee$. Further, since $M \prec_e M_\omega$ we have by Lemma D.2.4 that M_ω is also not short recursively saturated. □

This gives us the following interesting corollary, that any model satisfying collection has an elementary end extension which satisfies $\mathrm{Coll}\bigvee$ and which is tall.

Corollary D.2.12. *Every model M where $M \models \mathrm{PA}+\mathrm{Coll}$ has an elementary end extension N such that $N \models \mathrm{Coll}\bigvee$ and N is tall.*

Proof. Sketch We repeat the proof of the above, example, with the addition that when we build our end-extensions M_i we also build in the tallness criterion. We then take $N = M_\omega$ and we have the required elementary end extension. □

We thus know that there are models of Coll$\bigvee$ which are tall and not short recursively saturated.

Interestingly, this is not the only dependency between models of Coll$\bigvee$, tallness and short recursive saturation. We now prove that if a model satisfies Coll$\bigvee$ and is not tall, then it is short recursively saturated.

Lemma D.2.13. *Suppose $M \models \mathrm{PA} + \mathrm{Coll}\bigvee$ and M is not tall, then M is short recursively saturated.*

Proof. Since M satisfies PA we know it has definable Skolem terms We enumerate these terms $t_1(\bar{a}), t_2(\bar{a}), ...$ and suppose without loss of generality that $t_i(\bar{a}) < t_{i+1}(\bar{a})$ by taking $t'_{i+1}(\bar{a}) = \max\{t_i(\bar{a}) + 1, t_{i+1}(\bar{a})\}$. Since M is not tall, we know that $M = \lim_{n\in\mathbb{N}} t_n(\bar{a})$. Now suppose for contradiction that M is not short recursively saturated, i.e. there exist formulas θ_n such that $M \models \bigwedge_{n\in\mathbb{N}} \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a})$ and $M \models \forall x{<}t \bigvee_{n\in\mathbb{N}} \neg\theta_n(x, \bar{a})$. Then, for every $x < t$, there must be a least such $\theta(x, \bar{a})$ that is not satisifed. Formally, we can write this as:

$$M \models \forall x{<}t\exists y \bigvee_{n\in\mathbb{N}} [t_n(\bar{a}) \leqslant y \wedge y < t_{n+1}(\bar{a}) \wedge \bigwedge_{i<n} \theta_i(x, \bar{a}) \wedge \neg\theta_n(x, \bar{a})].$$

However, we know that:

$$M \not\models \exists z\forall x{<}t\exists y{<}z \bigvee_{n\in\mathbb{N}} [t_n(\bar{a}) \leqslant y \wedge y < t_{n+1}(\bar{a}) \wedge \bigwedge_{i<n} \theta_i(x, \bar{a}) \wedge \neg\theta_n(x, \bar{a})].$$

This is because if such a z existed, then for all $t_n(\bar{a}) < z$ we have that $M \vDash \forall x{<}t \bigvee_{i<n} \neg\theta_i(x, \bar{a})$. This contradicts our earlier supposition that $M \vDash \exists x{<}t \bigwedge_{i<n} \theta_i(x, \bar{a})$. Therefore, we have an instance where M fails Coll $\bigvee$ and therefore we have a contradiction and M must be short recursively saturated.

□

We now prove that if M is tall, then $M \vDash \text{Coll} \bigvee$.

Lemma D.2.14. *If $M \vDash \text{PA}$ is tall, then $M \vDash \text{Coll} \bigvee$.*

Proof. Suppose that $M \vDash \forall x{<}t \exists y \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$ and that M is tall. Then there is some w such that all Skolem terms $t_n(\bar{a}) < z$. Therefore, we have that $M \vDash \forall x{<}t \exists y{<}z \bigvee_{n\in\mathbb{N}} \theta_n(x, y, \bar{a})$ and so $M \vDash \text{Coll} \bigvee$. □

This allows us to show that $M \vDash \text{Coll} \bigvee$ if and only if it is tall or short recursively saturated by combining the previous lemmas.

Theorem D.2.15. *For any model $M \vDash \text{PA}$ we have that $M \vDash \text{Coll} \bigvee$ if and only if M is short recursively saturated or M is tall.*

Proof. Lemma 2.1.7 and lemma D.2.14 prove one direction of the theorem. For the other direction, if $M \vDash \text{Coll} \bigvee$, then if M is tall we are done, else M is not tall and we are done by lemma D.2.13. □

There is also an interesting connection between tallness and models of Coll $\bigwedge$. We prove below that all nonstandard models which satisfy Coll $\bigwedge$ are tall.

Lemma D.2.16. *If a model $M \vDash \text{PA} + \text{Coll} \bigwedge$ and $M \ncong \mathbb{N}$, then M is tall.*

Proof. We prove this using conditional proof. First suppose that t is a non-standard number. We know that:

$$M \vDash \forall x{<}t \exists y \bigwedge_{\ulcorner\theta\urcorner \in \mathbb{N}} [(n = x \wedge \exists v \theta(v)) \to \exists v{\leqslant}y \theta(v)].$$

This means that any formula with Gödel Number x, if the formula is standard and is satisfied, then there is some number larger than one of the constants that satisfies it. Thus, by Coll $\bigwedge$ we know:

$$M \vDash \exists z \forall x{<}t \exists y{<}z \bigwedge_{\ulcorner\theta\urcorner \in \mathbb{N}} [(n = x \wedge \exists v \theta(v)) \to \exists v{\leqslant}y \theta(v)].$$

In other words, for any standard formula which is satisfied, z is larger than one of the numbers which satisfies that formula. Thus M is tall. □

It is an open question as to what other results and interdependencies exist between the properties of: Coll $\bigwedge$, Coll $\bigvee$, tallness and short recursive saturation. It would, in particular, be interesting to explore whether a converse to Lemma D.2.16 exists and what connections exist between Coll $\bigwedge$ and short recursive saturation.

Question D.2.17. *Are there any further connections between* Coll $\bigwedge$, Coll $\bigvee$, *tallness and short recursive saturation?*

D.3 Second Order Collection

We shall now look at the second order variant of the collection schema. This axiom runs over, not just formulas, but also all subsets of a model. This is a highly interesting axiom as although only one countable model satisfies it, it is also a weak axiom in some instances, despite being second order.

D.3.1 The Second Order Collection Axiom

We start by defining what the second order collection axiom is.

Definition D.3.1 (Second Order Collection). *The second order collection axiom (*Coll^2*) is the sentence:*

$$\forall X \forall t [\forall x{<}t \exists y(\langle x, y\rangle \in X) \to \exists z \forall x{<}t \exists y{<}z(\langle x, y\rangle \in X)]$$

where X ranges over all subsets of the model M we are considering [10, Exercise 8.9].

Again, like first the first order collection schema, it is easy to intuit that $\mathbb{N}$ satisfies the second order collection axiom. To see this consider again that for a given set X we know that $\{x \in \mathbb{N} : x < t \text{ and } \exists y \langle x, y\rangle \in X\}$ is finite. Therefore there are finitely many y which will satisfy all the x in the above set, and so we can bound these y by some natural number larger than the maximum y. We shall see later that $\mathbb{N}$ is in fact the only countable set which does satisfy Coll^2.

D.3.2 Cardinals and Second Order Collection

In this subsection we will prove some interesting results about which models satisfy PA^- and Coll^2. Our first theorem in this section is that, surprisingly, any model of PA^- which is κ-like for a regular cardinal κ satisfies the second order collection axiom.

Theorem D.3.2. *Let κ be a regular cardinal and $M \models \mathrm{PA}^-$ where M is κ-like. Then $M \models \mathrm{Coll}^2$ [10, Exercise 8.9].*

Proof. Let $X \subseteq M$ and $t \in M$. Suppose, that $M \models \forall x{<}t \exists y \langle x, y\rangle \in X$ and, for contradiction, that hence:

$$M \models \forall z \exists x{<}t \forall y{<}z \langle x, y\rangle \notin X \ (\star).$$

For each $x < t \in M$ we fix a $y \in M$ such that $\langle x, y\rangle \in X$. Hence for all $z \in M$ there is $x < t$ such that $y > z$ (from $\star$). Then we have that $M = \bigcup\{\{z \in M : z \leqslant y\}\ x < t\}$. However this is a contradiction, as $\{z : z \leqslant y\} < \kappa$ and $\{x < t\} < \kappa$ since M is κ-like. Thus $M \models \mathrm{Coll}^2$. □

The converse to this theorem is also true, as we shall prove below.

Theorem D.3.3. *If $M \models \mathrm{PA}^- {+} \mathrm{Coll}^2$ and $|M| = \kappa$ where κ is cardinal, then M is κ-like [10, Exercise 8.9].*

Proof. Suppose $M \models \mathrm{PA}^- {+} \mathrm{Coll}^2$, $|M| = \kappa$ and suppose for contradiction that M is not κ-like. So there is some $a \in M$ such that $|\{x \in M : x < a\}| = \kappa$. Thus, we have a bijection $f : \{x \in M : x < a\} \to M$. We define the set $X = \{\langle x, y\rangle : x < a \text{ and } y = f(x)\}$. Then we know $M \models \forall x{<}a \exists y \langle x, y\rangle \in X$ and have that $M = \{f(x) : x < a\}$. Thus, by second order collection we have that $M \models \exists z \forall x{<}a \exists y{<}z \langle x, y\rangle \in X$. Therefore, we have that $M = \{f(x) < z : x < a\}$. This tells us that $z \notin M$ which is a contradiction, so thus M must be κ-like. □

Interestingly, this is not the case for singular cardinals. It turns that no model of cardinality κ, where κ is singular, satisfies second order collection. To prove this we first prove the following lemma about κ-like models.

Lemma D.3.4. *If $M \models \mathrm{PA}^-$, κ is a cardinal and M is κ-like, then for all cardinals $c < \kappa$ there is some $m \in M$ such that $|\{x \in M : y \leqslant m\}| \geqslant c$.*

Proof. Suppose that for some cardinal c we have that $|\{x \in M : x \leqslant m\}| < c$ for all $m \in M$. We then take some set $S \subseteq M$ of cardinality c. Therefore, by our supposition we know that S is cofinal in M. Thus, we have that $M = \bigcup_{s \in S}\{x \in M : x < s\}$. Hence, M has cardinality c which is a contradiction unless $c = \kappa$. □

This lemma allows us to prove that all models which have a cardinality of a singular cardinal cannot satisfy second order collection.

Theorem D.3.5. *If $M \models \mathrm{PA}^-$, κ is a singular cardinal and $|M| = \kappa$, then $M \not\models \mathrm{Coll}^2$ [10, Exercise 8.9].*

Proof. Suppose firstly that $M \models \mathrm{PA}^-$ is κ-like, for if it isn't, then by the contrapositive to Theorem D.3.3 it cannot satisfy second order collection and we are done. Now suppose that $\kappa = \sum_{1 \leqslant i < \lambda} \mu_i$, where $\lambda < \kappa$ and $\mu_i < \kappa$ for all i. Suppose without loss of generality that if $i < j$, then $mu_i < mu_j$. By the previous lemma, for each $i < \lambda$, we can take $x_i \in M$ such that $|\{a \in M : a \leqslant x_i\}| \geqslant \mu_i$. We then take $t \in M$ such that $|\{a \in M : a \leqslant t\}| \geqslant \lambda$. Thus, we have that $\{x_i : i < \kappa\}$ is cofinal in M, but that $\{a \in M : a \leqslant t\}$ is not. We can take a function $f : \{a \in M : a \leqslant t\} \to \{x_i : i < \lambda\}$. If we now consider $X = \{\langle x, y\rangle : x \leqslant t \text{ and } y = f(x)\}$, then we have that $M \models$

$\forall x{<}t\exists y\langle x, y\rangle \in X$. However, we know that $M \not\models \exists z \forall x{<}t\exists y{<}z\langle x, y\rangle \in X$, since $\mathrm{im}(f)$ is cofinal in M, but $\mathrm{dom}(f)$ is not. $\square$

Our next major result on second order collection is the following surprising theorem. We have that no nonstandard countable model of PA^- satisfies Coll^2 and thus, as $\mathbb{N} \models \mathrm{Coll}^2$, then only the standard countable model does so.

Theorem D.3.6. *There is exactly one countable model* $M \models \mathrm{PA}^-$ *such that* $M \models \mathrm{Coll}^2$ *[10, Exercise 8.9].*

Proof. By Example A.2.8 we know that $\mathbb{N}$ is ω-like and ω is a regular cardinal. It follows from Theorem D.3.2 that $\mathbb{N} \models \mathrm{Coll}^2$. It hence remains to show that there are no nonstandard countable models that satisfy Coll^2.

Suppose for contradiction that $M \models \mathrm{PA}^-$, $M \models \mathrm{Coll}^2$, M is countable and $M \not\cong \mathbb{N}$. We choose an (actually) infinite sequence $(a_0, a_1, a_2, ...)$ where $a_i < a_{i+1}$ for all $i \in N$ and $a_0 > \mathbb{N}$. This ensures that for every $x \in M$ there is some $i \in \mathbb{N}$ such that $a_i > x$. Now let $S \subseteq \mathbb{N}$. We now define the set

$$X_S = \{\langle x, y\rangle : \text{ for some } i \in \mathbb{N}, y > a_i \text{ and } \neg(p_i | x \Leftrightarrow i \in S)\}.$$

We shall use this to prove that M codes all sets $S \subseteq \mathbb{N}$ giving us a contradiction.

Suppose that S is not coded in M. Then, by Theorem A.1.2, S is not coded by any $x \in M$ where $x < a_0$. So for every $x < a_0$ there is some $i \in \mathbb{N}$ such that $\neg(p_i | x \Leftrightarrow i \in S)$. Hence for all such $x < a_0$ there is $y \in M$ where $\langle x, y\rangle \in X_S$. In other words: $M \models \forall x{<}a_0\exists y\langle x, y\rangle \in X_S$. Thus, by Coll^2 we have that:

$$M \vDash \exists z \forall x{<}a_0 \exists y{<}z \langle x, y\rangle \in X_S.$$

Without loss of generality take $z = a_n$ for some $n \in \mathbb{N}$ (this is because if $a_{n-1} < z < a_n$, then the statement holds for $z = a_n$ also). Hence for every $x \in M$ such that $x < a_0$ there is $y < a_n$ such that $\langle x, y\rangle \in X_S$. Expanding this, we get that for every $x \in M$ such that $x < a_0$ there is $y < a_n$ such that, for some $i \in \mathbb{N}$, $y > a_i$ and $\neg(p_i | x \Leftrightarrow i \in S)$. Therefore, since $a_i < y < a_n$ we must have that $i \in \{0, 1, 2, ..., n-1\}$. Hence, this tells us that for every $x \in M$ with $x < a_0$ there is some $i \in \{0, 1, 2, ..., n-1\}$ such that $\neg(p_i | x \Leftrightarrow i \in S)$. Thus we have that there is no $x \in M$ such that x codes $\{i \in \mathbb{N} : i < n$ and $i \in S\}$. However this is a contradiction with Theorem A.1.9, as this set is clearly finite.

Therefore we have that M codes all sets $S \subseteq N$. However there are 2^ω such subsets, and by hypothesis M is countable. This is a contradiction, so M cannot exist and the only countable model which satisfies the theorem is $\mathbb{N}$. □

D.3.3 Weakness of Coll^2

Given that only one countable model satisfies the axiom, we might form the opinion that this axiom is very strong. Surprisingly, in some cases, the axiom is instead rather weak. We now prove that for the class of Π_1-sentences of $\mathscr{L}_A$ that if PA$^-$ can prove any such sentence using Coll^2, then it can also prove it without Coll^2. Thus, surprisingly, second order collection is conservative over Π_1 sentences.

Theorem D.3.7. *If* $\mathrm{PA}^- + \mathrm{Coll}^2 \vdash \sigma$ *where* $\sigma \in \Pi_1$*, then* $\mathrm{PA}^- \vdash \sigma$*.*

Proof. Suppose that $\mathrm{PA}^- + \mathrm{Coll}^2 \vdash \sigma$ where $\sigma \in \Pi_1$ and suppose that $M \vDash \mathrm{PA}^-$. We build an end extension M_1 of M and then repeat this process on a model M_i to build an end extension of it M_{i+1}. We take the union of these at each ordinal and arrive at M_{ω_1} as the union of the previous models which is the first uncountable model. We have stated that ω_1 is regular. We know M_{ω_1} is ω_1-like, since it is the first uncountable model arrived at. Hence by Theorem D.3.2 we know that $M_{\omega_1} \vDash \mathrm{Coll}^2$. Therefore $M_{\omega_1} \vDash \sigma$ and thus, since σ is Π_1 we have that $M \vDash \sigma$. Therefore, by completeness, $\mathrm{PA}^- \vdash \sigma$. □

The proof of this example also shows that it is easy to build models which satisfy second order collection. We prove this explicitly below. This provides a further example of how second order collection can also be weak, as it can have a large amount of models.

Theorem D.3.8. *For all theories* T *where* $T \supseteq \mathrm{PA}$ *there is a model of* T *which satisfies second order collection.*

Proof. Let $T \supseteq \mathrm{PA}$ be a theory. Thus T has a model M_0, which without loss of generality is countable by the Löwenheim-Skolem Theorems. Since $M_0 \vDash \mathrm{PA}$ we know by the MacDowell-Specker Theorem, Theorem A.1.1, that M_0 has an elementary end extension M_1. For each model M_i, where $i \geqslant 1$ we build an elementary end extension M_{i+1}. We can do this since $M_i \succ_e M_{i-1}$ and $M_{i-1} \vDash \mathrm{PA}$. At each limit ordinal λ we build $M_\lambda = \bigcup_{i<\lambda} M_i$. In this way we construct a chain of models, and we stop once we reach M_{ω_1} which is the first uncountable model. We know M_{ω_1} is ω_1-like, since $M_{\omega_1} = \bigcup_{i<\omega_1} M_i$, so otherwise we have that M_i is uncountable for some $i < \omega_1$ which is a contradiction. Thus, since ω_1 is regular, we know that $M_{\omega_1} \models \mathrm{Coll}^2$ by Theorem D.3.2. Thus, since $M_{\omega_1} \succ_e M_0$ we know that $M_{\omega_1} \models T$ and we are done. $\square$

Appendix E

Satisfaction Classes

In this chapter we examine the theory of satisfaction classes. We begin with an exploration of Tarski's definition of truth for arithmetic and then show how this is developed to define satisfaction classes. We then provide two of the key theorems within this area, as well as the formal tools to prove them. Finally, we explore modifications of the theory of satisfaction classes within the literature, as well as two suggestions of our own.

E.1 Tarski and Truth

Whilst truth is an intuitive concept, a full logical description of truth is a difficult philosophical project. One of the main approaches to doing this, and the one which bases truth in mathematics and satisfaction classes, is Tarski's work. In this section we shall give an overview of his work and in its main features. These main features, to be explored in the following subsections, are:

a) An expression is true if and only if its meaning holds.

b) Through an assignment of variables we can talk about the truth of formulas.

c) The truth of a connected sentence can be understood by the truth of the connectives.

d) Truth is not a definable concept.

Throughout this section we shall use Tr as a new predicate of the language $\mathscr{L}_A \cup \{Tr\}$ where we intend to write $Tr(X)$ to mean X is true.

E.1.1 Truth of Sentences in Terms of Meaning

In this subsection we look at feature a), that we understand the truth of an expression based on whether its meaning is the case or not. The canonical example of this is: 'snow is white' is true if and only if snow is white. Symbolically, for a sentence S, we write $Tr(\ulcorner S \urcorner) \iff S$. This is intuitively attractive, if S holds, then we want to be able to say that the expression stating S is true. Similarly, if an expression stating S is true, then we want S to hold.

Whilst this is obvious in English, this hides a deep idea about truth. In English we use snow is white in two separate ways. We use it to represent the concept of the whiteness of snow and to represent the proposition 'snow is white'. From Tarski [20, Section 1] we have that when talking about truth we talk about propositions, rather than concepts, but the truth of a proposition is analysed in terms of the attached concept.

For an example suppose we write $D(l)$ to write the proposition that lead is dense. We do not have that the concept that lead is dense is true. However, because lead is dense holds, we write that $D(l)$ is true. Thus, when analysing the truth of arithmetical statements in our language $\mathscr{L}$ we talk of the truth of the Gödel Number of these sentences. We hence write that a model M with a truth predicate Tr satisfies:

$$(M, Tr) \vDash Tr(\ulcorner \sigma \urcorner) \text{ if and only if } M \vDash \sigma.$$

A simple example is that we write $(\mathbb{N}, Tr) \vDash Tr(\ulcorner 0 \leqslant 1 \urcorner)$ because $\mathbb{N} \vDash 0 \leqslant 1$.

This is a subtle point, which we shall often pass over for ease of communication, but is a powerful idea.

E.1.2 Truth of Formulas

Whilst in the above section we have an analysis for the truth of sentences, this leads to difficulties in analysing the truth of formulas. To say that $(M, Tr) \vDash Tr(\ulcorner 1 + 1 = 2 \urcorner)$, we require a constant symbol for 2 in our language $\mathscr{L}_{Tr} \supseteq \mathscr{L}_A$. Whilst this is relatively innocuous in this case, to talk about the truth of formulas involving any (potentially nonstandard) number a, for example $(M, Tr) \vDash Tr(\ulcorner a = 0 \vee a > 0 \urcorner)$, we require a constant symbol for every single number, and our language is no longer finite. This is a direct consequence of the insights from the previous subsection, that truth is a predicate of Gödel Numbers of sentences, rather than the meaning of those sentences.

There are two ways of getting around this problem, each with their own advantages and disadvantages. The first option is to perform a syntactic operation on formulas which substitutes free variables in a formula by canonical terms, sentences which define a unique number in the model.

By $clterm(a)$ we denote the PA-definable canonical term of a, which we now define below. We use $\frown$ to denote the concatenation symbol and introduce the notation $\llcorner g \lrcorner$ which is the formula taking a Gödel-number g and mapping it to the corresponding formula $\theta(\bar{x})$ of $\mathscr{L}_A$, such that $\ulcorner \theta(\bar{x}) \urcorner = g$. We then define the canonical term of a number $a \in M$ in the following manner:

Definition E.1.1.

$$clterm(0) = \ulcorner 0 \urcorner$$

$$clterm(n+1) = \ulcorner (\frown \llcorner clterm(n) \lrcorner \frown) \urcorner, \textit{for } n \geqslant 0 \ [\textit{12}, Exercise 9.11].$$

With canonical terms in place, we can use the PA-definable formula $subs(\ulcorner \theta \urcorner, x, a) = \ulcorner \psi \urcorner$ to write that ψ is obtained from the (potentially non-standard) formula $\theta(x)$ by replacing every free occurrence of the variable x by the the canonical term of $a, clterm(a)$ as can be seen in Exercise 9.5 of *Models of Peano ARithmetic* [12]. We then get that:

$$M \vDash \theta(a) \text{ if and only if } (M, Tr) \vDash Tr(subs(\ulcorner \theta \urcorner, x, a)).$$

We can then expand this method to work for formulas with more than one free variable. The formula *subs* can be defined in various ways, by specifying the axioms that it must satisfy. Whilst this is do-able in PA the disadvantage of this method is that it does not work in all theories.

The other option, and the one that we shall take, is to define truth over formulas and a corresponding coding of an assignment of free variables to elements of the model. Our assignment is a function $f : S \to M$ where S is a finite set of variables from $\mathscr{L}_A$. We then code this function with the help of the β-function by $a \in M$ such that $(a)_i = f(x_i)$. We make Tr a two-place predicate and we write that a model M with Tr satisfies:

$$(M, Tr) \vDash Tr(\ulcorner \theta(\bar{x}) \urcorner, a) \text{ if and only if } M \vDash \theta(\bar{c}),$$

where $\bar{c} \in M$ is such that $\bar{c}$ is mapped to by the function coded by a.

This is not easy to setup and requires a lot of technical working. The details of this procedure can be found in Chapter 9 of *Models of Peano Arithmetic* [12].

With Tr defined in such a way as to make this work, truth about (atomic) formulas can now be expressed with only finitely many constant symbols. For example, we can now write for $M \vDash$ PA that $(M, Tr) \vDash Tr(\ulcorner 1 + 1 = x \urcorner, b)$, where b codes the assignment of x to 2, since $M \vDash 1 + 1 = 2$.

With this system in place we now need to clarify how we define truth for sentences using our new two-place truth predicate we can do this very easily.

Definition E.1.2. *We have for any $\mathscr{L}_A$-sentence σ which is free from connectives that $(M, Tr) \vDash Tr(\ulcorner \sigma \urcorner, a)$ if and only if $M \vDash \sigma$ for any assignment a.*

With this analysis in place we have offered a truth definition of atomic formulas and sentences, but a definition for formulas involving logical connectives and quantifiers is still missing. We now look at this in the next subsection.

E.1.3 Truth of Connected Sentences

With a definition of truth for the (Gödel Number of) atomic formulas in place, we are able to define truth for non-atomic formulas as well, by an inductive definition over the connectives. To do this, we describe some new notation for representing a new assignment based on a previous one.

Notation E.1.3. *Suppose we have a formula $\theta(x, \bar{y})$ and an assignment a that maps the free variables $x, \bar{y}$ to corresponding canonical terms of numbers. We then write $a[^x/_b]$ to denote the assignment that maps the free variables $\bar{y}$ to the constants a and maps the free variable x to b.*

We take $TrAt$ as our truth predicate for atomic formulas and use the $\mathscr{L}_A$-definable predicate $At(\theta(\bar{x}))$ to represent that $\theta(\bar{x})$ is an atomic formula. We then define truth for all formulas using the following definition.

Definition E.1.4 (Tarski's Definition of Truth)**.** *We expand our language from $\mathscr{L}_A$ to $\mathscr{L}_{Tr} = \mathscr{L}_A \cup \{Tr\}$ where Tr is a new two-place predicate. We have the following axioms that govern the behaviour of Tr. For all $\mathscr{L}_A$-formulas $\theta(\bar{x})$, $\alpha(\bar{x})$ and $\beta(\bar{x})$ we have:*

Axiom 1: $At(\theta(\bar{x})) \rightarrow [TrAt(\ulcorner\theta(\bar{x})\urcorner, a) \leftrightarrow Tr(\ulcorner\theta(\bar{x})\urcorner, a)]$.

Axiom 2: $Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \wedge Tr(\ulcorner\beta(\bar{x})\urcorner, a) \leftrightarrow Tr(\ulcorner\alpha(x) \wedge \beta(x)\urcorner, a)$.

Axiom 3: $Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \vee Tr(\ulcorner\beta(\bar{x})\urcorner, a) \leftrightarrow Tr(\ulcorner\alpha(x) \vee \beta(x)\urcorner, a)$.

Axiom 4: $Tr(\ulcorner\neg\theta(\bar{x})\urcorner, a) \leftrightarrow \neg Tr(\ulcorner\theta(\bar{x})\urcorner, a)$.

Axiom 5: $Tr(\ulcorner\exists x\theta(x,\bar{y})\urcorner, a) \leftrightarrow \exists b Tr(\ulcorner\theta(x,\bar{y})\urcorner, a[^x/_b])$.

Axiom 6: $Tr(\ulcorner\forall x\theta(x,\bar{y})\urcorner, a) \leftrightarrow \forall b Tr(\ulcorner\theta(x,\bar{y})\urcorner, a[^x/_b])$ $[20, Definition 22]$.

Whilst no axiom is given, this also gives us the tools to understand the truth of a sentence involving a conditional arrow. For completeness, we explain this below, where we understand $Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \rightarrow Tr(\ulcorner\beta(\bar{x})\urcorner, a)$ to mean $\neg Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \vee Tr(\ulcorner\beta(\bar{x})\urcorner, a)$.

Lemma E.1.5. *Suppose $M \vDash$ PA and Tr is a truth predicate for M as above. Then for all $\mathscr{L}_A$-formulas $\alpha(\bar{x})$ and $\beta(\bar{x})$:*

$$(M, Tr) \vDash Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \rightarrow Tr(\ulcorner\beta(\bar{x})\urcorner, a) \leftrightarrow Tr(\ulcorner\alpha(\bar{x}) \rightarrow \beta(\bar{x})\urcorner, a).$$

Proof. Suppose firstly that $(M, Tr) \vDash Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \rightarrow Tr(\ulcorner\beta(\bar{x})\urcorner, a)$. Then we have that:

$$(M, Tr) \vDash \neg Tr(\ulcorner\alpha(\bar{x})\urcorner, a) \vee Tr(\ulcorner\beta(\bar{x})\urcorner, a).$$

Therefore, by the Tarski Axioms 3 and 4 we have that $(M, Tr) \vDash Tr(\ulcorner\neg\alpha(\bar{x}) \vee \beta(\bar{x})\urcorner, a)$ and hence we can rewrite this as:

$$(M, Tr) \vDash Tr(\ulcorner\alpha(\bar{x}) \rightarrow \beta(\bar{x})\urcorner, a).$$

Now suppose that $(M, Tr) \vDash Tr(\ulcorner\alpha(\bar{x}) \rightarrow \beta(\bar{x})\urcorner, a)$. Then by the laws of logic $(M, Tr) \vDash Tr(\ulcorner\neg\alpha(\bar{x}) \vee \beta(\bar{x})\urcorner, a)$. Thus, by applying the Tarski Axioms

3 and 4 we get that:

$$(M, Tr) \vDash \neg Tr(\ulcorner \alpha(\bar{x}) \urcorner, a) \vee Tr(\ulcorner \beta(\bar{x}) \urcorner, a)$$

and thus, by definition, $(M, Tr) \vDash Tr(\ulcorner \alpha(\bar{x}) \urcorner, a) \rightarrow Tr(\ulcorner \beta(\bar{x}) \urcorner, a)$. □

We thus have a definition from Tarski that defines truth for all standard $\mathscr{L}_A$-formulas and sentences. The following example shows how this works in practice.

Example E.1.6. *Suppose that* $(M, Tr) \vDash$ PA+ *Axioms 1-6 for* Tr*. Then* $(M, Tr) \vDash Tr(\ulcorner \exists x(2 + x = 5) \wedge \neg \forall y(y < 0) \urcorner, a)$ *for any assignment* a*.*

Proof. We know $(M, Tr) \vDash At(2 + x = 5)$ and, for any assignment a, that:

$$(M, Tr) \vDash TrAt(\ulcorner 2 + x = 5 \urcorner, a[^x/_3]),$$

since $M \vDash 2 + 3 = 5$. Therefore, by Tarski Axiom 1, we know:

$$(M, Tr) \vDash Tr(\ulcorner 2 + x = 5 \urcorner, a[^x/_3]).$$

Therefore $(M, Tr) \vDash \exists b[Tr(\ulcorner 2 + x = 5 \urcorner, a[^x/_b])]$. We now can now apply Tarski Axiom 5 to get that $(M, Tr) \vDash Tr(\ulcorner \exists x[2 + x = 5] \urcorner, a)$.

Similarly, we know $(M, Tr) \vDash At(y < 0)$ and for any assignment a, that $(M, Tr) \vDash \neg TrAt(\ulcorner y < 0 \urcorner, a)$. So therefore, by Tarski Axiom 1, we have that:

$$(M, Tr) \vDash \neg Tr(\ulcorner y < 0 \urcorner, a).$$

Suppose now for some assignment a, that $(M, Tr) \vDash \forall b Tr(\ulcorner y < 0 \urcorner, a[{}^{y}/_{b}])$. This immediately contradicts with our last derivation, so thus:

$$(M, Tr) \vDash \neg \forall b Tr(\ulcorner y < 0 \urcorner, a[{}^{y}/_{b}]).$$

We can now apply the Tarski Axioms 6 and 4 to get that:

$$(M, Tr) \vDash Tr(\ulcorner \neg \forall y [y < 0] \urcorner, a).$$

Therefore, we have that:

$$(M, Tr) \vDash Tr(\ulcorner \exists x [2 + x = 5] \urcorner, a) \wedge Tr(\ulcorner \neg \forall y [y < 0] \urcorner, a).$$

Finally, we can now apply Tarski Axiom 2 to get that:

$$(M, Tr) \vDash Tr(\ulcorner \exists x (2 + x = 5) \wedge \neg \forall y (y < 0) \urcorner, a)$$

and we see that this holds for all assignments a. □

This theorem from Tarski shows that we did not pick the above sentence specially, and in fact Tarski's definition of truth is satisfactory for our original aim.

Theorem E.1.7. *For all standard $\mathscr{L}_A$-formula $\theta(\bar{x})$ there is a suitable assignment a such that $(M, Tr) \vDash Tr(\ulcorner \theta(\bar{x}) \urcorner, a) \leftrightarrow \theta(\bar{c})$ [20, Section 5, Theorem 2].*

For a proof of this theorem we refer the reader to Tarski's *Concept of*

Truth in Formalised Languages [20].

E.1.4 Undefinability of Truth

We have hence seen that from Tarski's definition of truth we can define what it means for a standard formula to be true or false. We have set this definition up to be an extension of the language $\mathscr{L}_A$. The natural question that arises, given this extension, is whether we really need to extend our language. Is it possible to define a formula in $\mathscr{L}_A$ which satisfies the Tarski axioms? It is a well-known theorem that some models of PA contain an element which can define truth for standard sentences of $\mathscr{L}_A$. We prove this theorem now below.

Theorem E.1.8. *There exists models $M \models \mathrm{PA}$ which contain an element $a \in M$ such that $M \models (a)_{\ulcorner\sigma\urcorner} = 1$ if and only if $M \models \sigma$ for all standard $\mathscr{L}_A$-sentences σ. In other words, a codes a sequence such that the (Gödel-number of the) 'σth' position of a is 1 if and only if the sentence σ is true in M.*

Proof. Let $M \models \mathrm{PA}$ such that M is recursively saturated. We enumerate all standard $\mathscr{L}_A$-sentences as $\sigma_1, \sigma_2, \sigma_3, \ldots$ and consider the type $\{\phi_i(x) : i \in \mathbb{N}\}$. We take $\phi_i(x)$ to be the formula:

$$\sigma_i \leftrightarrow (x)_{\ulcorner\sigma_i\urcorner} = 1.$$

We know that all models $M \models \mathrm{PA}$ satisfy $M \models \exists x \bigwedge_{i<n} \phi_i(x)$ for any $n \in \mathbb{N}$. To see this fix n and take $a \in M$ that codes a finite sequence of length:

$$l = max\{\ulcorner\sigma_i\urcorner : i < n\}.$$

We further specify that $(a)_i = 0$ for $i \leqslant l$, unless $M \vDash \sigma_i$, in which case $(a)_{\ulcorner \sigma_i \urcorner} = 1$. We are able to do this using the Gödel-beta lemma, see Section A.1.2. We thus have that $M \vDash \exists x \bigwedge_{i<n} \phi_i(x)$ for all $n \in \mathbb{N}$, so since M is recursively saturated:

$$M \vDash \exists x \bigwedge_{n \in \mathbb{N}} \phi_n(x).$$

□

Given this result, one may be tempted to answer our question in the affirmative, that it is possible to define truth in $\mathscr{L}_A$. Tarksi proved that this is in fact not the case, and truth cannot be defined for $\mathscr{L}_A$ within $\mathscr{L}_A$. To prove this theorem, we first state and prove a slightly stronger version of the Diagonal Lemma, by adapting the proof found in *Models of Peano Arithmetic* [12, Lemma 3.13].

Lemma E.1.9 (Diagonal Lemma)**.** *Let $T \supseteq$ PA be a theory in the language $\mathscr{L}_A$ and $\theta(x, \bar{z})$ be an $\mathscr{L}_A$-formula. Then there is an $\mathscr{L}_A$-formula $G(\bar{z})$ such that $T \vdash \forall \bar{z}[G(\bar{z}) \leftrightarrow \theta(\ulcorner G(\bar{z}) \urcorner, \bar{z})]$.*

Proof. We begin by considering the function:

$$diag(n, \bar{z}) = \begin{cases} \ulcorner \forall y(y = n \to \sigma(y, \bar{z}) \urcorner & \text{if } n = \ulcorner \sigma(x, \bar{z}) \urcorner \text{ is the Gödel-number of an } \mathscr{L}_A\text{-formula} \\ 0 & \text{otherwise} \end{cases}$$

We then define the formula $\delta(x, y, \bar{z})$ which holds if and only if:

$$T \vdash diag(x, \bar{z}) = y.$$

We then define $\psi(x, \bar{z})$ as holding if and only if $\forall t[\delta(x, t, \bar{z}) \to \theta(x, \bar{z})]$. We now take $n = \ulcorner \psi(x, \bar{z}) \urcorner$. Then, we have our formula:

$$G(\bar{z}) \text{ is } \forall y[y = n \to \psi(y, \bar{z})].$$

Therefore, we see that $T \vdash \forall \bar{z}[G(\bar{z}) \leftrightarrow \psi(n, \bar{z})]$. Thus, we have that:

$$T \vdash \forall \bar{z}[G(\bar{z}) \leftrightarrow \forall t(\delta(n, t, \bar{z}) \to \theta(t, \bar{z}).$$

Hence:

$$T \vdash \forall \bar{z}[\forall t(\delta(n, t, \bar{z}) \leftrightarrow t = \ulcorner G(\bar{z}) \urcorner)].$$

Finally, we therefore have that $T \vdash \forall \bar{z}[G(\bar{z}) \leftrightarrow \theta(\ulcorner G(\bar{z}) \urcorner, \bar{z})]$. □

With this Diagonal Lemma in place it is easy to prove Tarski's Theorem on the undefinability of truth.

Theorem E.1.10 (Tarski's Theorem). *Suppose $M \models$ PA and $\bar{a} \in M$. There is no $\theta(x, \bar{a})$ definable in $\mathscr{L}_A \cup \{\bar{a}\}$ such that for all standard sentences $\sigma(\bar{a})$: $M \models \theta(\ulcorner \sigma \urcorner, \bar{a}) \leftrightarrow \sigma$ [20, 5, Theorem 1].*

Proof. Suppose $\bar{a} \in M$ and $\theta(x, \bar{a})$ is an $\mathscr{L}_A \cup \{\bar{a}\}$-formula as above. Then we know by the Diagonal Lemma that there is an $\mathscr{L}_A \cup \{\bar{a}\}$-sentence $G(\bar{a})$ such that:

$$(M, \bar{a}) \models G(\bar{a}) \leftrightarrow \neg\theta(\ulcorner G(\bar{a}) \urcorner, \bar{a}).$$

Thus, if $(M, \bar{a}) \models G(\bar{a})$, then by the definition of $\theta(x, \bar{a})$ we have that $(M, \bar{a}) \models \neg G(\bar{a})$ which is a contradiction. But if $(M, \bar{a}) \models \neg G(\bar{a})$, then by $\theta(x, \bar{a})$ we have that $(M, \bar{a}) \vDash G(\bar{a})$ which is similarly a contradiction. So no such formula $\theta(x, \bar{a})$ can exist. $\square$

Tarski's Theorem hence tells us that the Tr in Tarski's definition of truth is a new predicate and cannot be defined by any $\mathscr{L}_A$ formula. We hence distinguish notationally between the two languages by: $\mathscr{L}_A$ and $\mathscr{L}_{Tr}$, where $\mathscr{L}_{Tr} = \mathscr{L}_A \cup \{Tr\}$ is the language with a truth predicate Tr.

E.1.5 Tarski and Deflationary Truth

It is interesting to consider to what extent Tarski's account of truth is a deflationary account of truth. There are two ways to approach this, one can discuss whether Tarski's philosophical approach to truth is a deflationist conception of truth or whether Tarski's logical treatment of truth, as explicated above, is a deflationist one in the sense of Definition 5.2. The philosophical approach has been widely discussed in the literature and the reader is referred to Schantz's *Was Tarski a Deflationist?* [18] for an analysis of this. We shall focus on the latter question of whether Tarski's logical treatment of truth is deflationist in the sense of Definition 5.2, i.e. whether it satisfies the equivalence schema for all considered sentences and whether every considered model of PA has an expansion to this truth predicate.

Ketland comes to the conclusion that Tarski's definition of truth is not a deflationist account. He demonstrates [13, Page 81] that this theory is not conservative over PA, and thus not deflationist in either our sense or his own as per Definition C.1. However, these findings are based on taking the background theory in which induction runs over formulas containing the predicate Tr. We will not consider this yet.

Instead, we find that Tarski's definition can be deflationist. We began Tarski's account by explicating Tarski's aim that $(M, Tr) \vDash Tr(\ulcorner\sigma\urcorner)$ if and only if $M \vDash \sigma$. In Theorem E.1.7 we saw that Tarski's definition of truth satisfies this for all standard sentences of $\mathscr{L}_A$. Further, Tarski proves [20, Section 5, Theorem 3] that every model has an expansion to his truth predicate Tr. Therefore, if one considers standard sentences of the language of arithmetic, Tarski's definition of truth is certainly a deflationist account in the sense of our definition.

We would like to consider, however, sentences of arithmetic of nonstandard length, in the language ${}^{\star}\mathscr{L}_A$. These sentences can be built from a nonstandard number of connectives and are, from the perspective of a nonstandard model of PA, as important as standard sentences. Were Tr capable of being defined by an $\mathscr{L}_A$-formula then this would be easy, as we could use induction on formulas involving Tr. However, from Tarski's theorem we cannot do this. In the following section we shall define truth for nonstandard sentences of arithmetic using satisfaction classes. These are mathematical structures which establish that the Tarski conditions hold for nonstandard sentences as well and have some highly interesting properties.

E.2 Satisfaction Classes

The theory of satisfaction classes has been developed to define truth for nonstandard sentences. Tarksi developed a definition of truth that works for all standard sentences, but this does not tell us anything about nonstandard sentences. This leads to an interesting and rich theory with many open problems.

Throughout this chapter we now consider models $M \vDash \mathrm{PA}$ where M is nonstandard. Whilst Tarski's axioms for truth offer a definition of truth for standard sentences of M, they do not provide one for nonstandard sentences of M. There are nonstandard elements $t \in M$ such that $M \vDash \llcorner t \lrcorner$ "is a formula $\theta(\bar{x})$". This formula is nonstandard, so contains ω-many connectives, so applying Tarski's definition of truth does not tell us the truth of this formula with the required assignment. For example, the sentence $\bigwedge_{x<t} x = x$ for nonstandard $t \in M$ is intuitively true, but applying Tarski's definition requires us to evaluate a nonstandard number of conjuncts, which we cannot induct over.

To evaluate the truth of these nonstandard formulas of M we use the notion of a satisfaction class. These structures, based on the Tarski axioms, are a natural way of strengthening the axioms to be applicable to sentences of nonstandard length. To do this, we first need to introduce the following notation, which determines whether a nonstandard number $t \in M$ is the Gödel-number of a formula as understood by a nonstandard model M.

Notation E.2.1. *We write $M \vDash form(t)$ if and only if t is the Gödel number of a formula in the sense of M. We write $M \vDash term(t)$ if and only if t is the Gödel number of a canonical term in M. We write $M \vDash val(t) = x$ if and only if $M \vDash clterm(x) = t$.*

The following lemma tells us that we are free to introduce this notation.

Lemma E.2.2. *The formulas $form(x), term(x)$ and $val(x)$ can all be defined* *in* PA *[12, Section 9.2].*

With this formula in place we are now in place to define satisfaction classes.

Definition E.2.3 (Full Satisfaction Class)**.** *We expand our language from* $\mathscr{L}_A$ *to* $\mathscr{L}_S = \mathscr{L}_A \cup \{S\}$*. A satisfaction class is a two-place predicate* S *over* M *such that* $(M, S) \vDash S(\ulcorner\theta\urcorner, a)$ *only if* $M \vDash form(\ulcorner\theta\urcorner)$*. If* $M \vDash form(\ulcorner\theta\urcorner)$ *and* $a \in M$*, then* $(M, S) \vDash S(\ulcorner\theta\urcorner, a)$ *if and only if* (M, S) *satisfies one of the following conditions:*

Tarski Condition 1) $\exists t, s[term(t) \wedge term(s) \wedge \ulcorner\theta\urcorner = \ulcorner(t = s)\urcorner \wedge val(t, a) = val(s, a)]$.

Tarski Condition 2) $\exists t, s[term(t) \wedge term(s) \wedge \ulcorner\theta\urcorner = \ulcorner(t < s)\urcorner \wedge val(t, a) < val(s, a)]$.

Tarski Condition 3) $\exists\alpha, \beta[form(\alpha) \wedge form(\beta) \wedge \ulcorner\theta\urcorner = \ulcorner(\llcorner\alpha\lrcorner \wedge \llcorner\beta\lrcorner)\urcorner \wedge (S(\alpha, a) \wedge S(\beta, a))]$.

Tarski Condition 4) $\exists\alpha, \beta[form(\alpha) \wedge form(\beta) \wedge \ulcorner\theta\urcorner = \ulcorner(\llcorner\alpha\lrcorner \vee \llcorner\beta\lrcorner)\urcorner \wedge (S(\alpha, a) \vee S(\beta, a))]$.

Tarski Condition 5) $\exists\psi[form(\psi) \wedge \ulcorner\theta\urcorner = \ulcorner\neg\llcorner\psi\lrcorner\urcorner \wedge \neg S(\psi, a)$.

Tarski Condition 6) $\exists i, \psi[form(\psi) \wedge \ulcorner\theta\urcorner = \ulcorner\exists x\llcorner\psi\lrcorner\urcorner \wedge \exists b S(\psi, a[^x/_b])$.

Tarski Condition 7) $\exists i, \psi[form(\psi) \wedge \ulcorner\theta\urcorner = \ulcorner\forall x\llcorner\psi\lrcorner\urcorner \wedge \forall b S(\psi, a[^x/_b])$.

[12, Definition 15.1]

A satisfaction class, therefore, behaves like the Tarskian truth predicate, with the requirement that it ranges over sentences of ${}^*\mathscr{L}_A(M)$ and not just $\mathscr{L}_A$.

This is not the only type of satisfaction class we can consider, although it shall be the main one that we do so. We can also consider a partial satisfaction class, that is defined only on some formulas.

Definition E.2.4 (Partial Satisfaction Class)**.** *A* Γ*-satisfaction class, where* $\Gamma \subseteq \{\theta(\bar{x}) : M \vDash form(\theta(\bar{x}))\}$*, is a satisfaction class for all formulas in* Γ*. If we introduce the predicate* $\Gamma(x)$ *such that* $M \vDash \Gamma(x)$ *if and only if* $x \in \Gamma$*, then we can write this more formally as:* S_Γ *is a* Γ*-satisfaction class for* M *if* S_Γ *is a two-place predicate over* M *such that* $M \vDash S_\Gamma(x, a)$ *if and only if* $M \vDash \Gamma(x)$*. If* $x \in \Gamma$ *and* $a \in M$*, then* $(M, S_\Gamma) \vDash S_\Gamma(x, a)$ *if and only if* (M, S_Γ) *satisfies one of the Tarski Conditions 1-7.*

A partial satisfaction class S *for* M *is a* Γ*-satisfaction class for some* Γ *which is a subset of* $\mathscr{L}$ *formulas [12, Definition 15.1].*

We define the connective $\rightarrow$ for satisfaction classes in the same manner as we did for the Tarskian truth predicate. We say that $S(\alpha, a) \rightarrow S(\beta, a) \leftrightarrow \neg S(\alpha, a) \vee S(\beta, a)$. It is a simple exercise to follow the proof of Lemma E.1.5 to show this holds if and only if $S(\alpha \rightarrow \beta, a)$.

These definitions are very general and thus lead to three interesting questions:

1. Does every model have a satisfaction class?

2. Of those models which do, are the satisfaction classes unique?

3. Is there a procedure to develop a satisfaction class for a model?

We shall answer the first two questions in this section and the third in later sections. First, we shall develop the notation which will be used to deal with satisfaction classes.

Notation E.2.5. *We write* (M, S) *to denote an* $\mathscr{L}_S$*-structure such that* $M \models \mathrm{PA}$ *and* S *is a full satisfaction class over* M*. We shall sometimes write* $S \models \sigma$ *to denote that* $(M, S) \models S(\sigma)$*.*

Before answering these three questions, we should first test whether a satisfaction class is fit for purpose. We prove below, as satisfaction classes follows Tarski's definition, that a satisfaction class defines truth for all standard formulas.

Theorem E.2.6. *For all standard* $\mathscr{L}_A$*-formulas* $\theta(\bar{x})$ *and constants* $c \in M$ *we have that there is* $a \in M$ *such that* $(M, S) \models S(\ulcorner \theta(\bar{x}) \urcorner, a) \leftrightarrow \theta(\bar{c})$*.*

Proof Sketch. Since $\theta(\bar{x})$ is standard we know it contains only an actually finite number of connectives. Thus, we can repeatedly apply the Tarski Conditions 1-7 on the subformulae of $\theta(\bar{x})$ to get that for a suitable assignment $a \in M$ we have that $(M, S) \models S(\ulcorner \theta(\bar{x}) \urcorner, a)$. If $(M, S) \models S(\ulcorner \theta(\bar{x}) \urcorner, a)$, then we can repeatedly use the Tarski Conditions 1-7, since $\theta(\bar{x})$ has a finite number of connectives, to get that that the satisfaction class makes all atomic subformulas of $\theta(\bar{x})$ true and hence by definition these atomic subformulas hold in M when given the constant $\bar{c}$. Thus, $(M, S) \models \theta(\bar{c})$. $\square$

This leads to the corollary, by Tarski's Theorem, there is no $\mathscr{L}_A$ formula which defines a satisfaction class S.

Corollary E.2.7. *There is no $\mathscr{L}_A$ formula which captures S.*

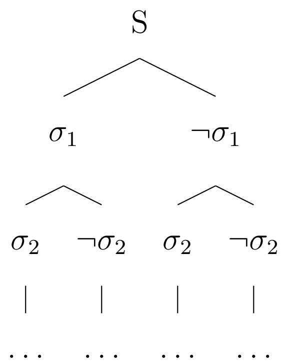
Proof. If there were an $\mathscr{L}_A$ formula $\theta(\bar{z})$ such that for all formulas $\phi(\bar{x})$ and assignments a, $(M, S) \models S(\ulcorner\phi(\bar{x})\urcorner, a) \leftrightarrow \theta(\ulcorner\phi(\bar{x})\urcorner, a, \bar{z})$, then by Theorem E.2.6 we have that $(M, Tr) \models Tr(\ulcorner\phi(\bar{x})\urcorner, a) \leftrightarrow \theta(\ulcorner\phi(\bar{x})\urcorner, a, \bar{z})$ which by Tarski's Theorem (Theorem E.1.10) is a contradiction. □

Whilst we have seen that satisfaction classes make every true standard formula true, this property does not hold for sentences of nonstandard length. Thus satisfaction classes can make intuitively true sentences false and intuitively false sentences true. The informal reason for this is because a satisfaction class can only consider a natural number of connectives and thus approximates a sentence to be true or false based on the finite part of the sentence that it can see. We prove an example of this in Example E.5.1.

That there are discrepancies between our intuitions of truth and what some satisfaction classes say are true leads to an answer to our second question. The intuition arises that one model can have many satisfaction classes, some of which contradict each other for truth on nonstandard sentences. In fact, it is the case that a countable model has $2^{\aleph_0}$ satisfaction classes. We prove this theorem below.

Theorem E.2.8. *A countable model M has $2^{\aleph_0}$ satisfaction classes [14].*

Proof. We can consider a satisfaction class as an infinite path through an infinite binary tree, where each node is σ or $\neg\sigma$ for an $\mathscr{L}_A$-sentence σ. Diagrammatically, we can picture this as:



We see that the maximum number of satisfaction classes a countable model can contain is clearly $2^{\aleph_0}$ since there are only countably many $\mathscr{L}_A$-sentences in a countable model and thus $2^{\aleph_0}$-many infinite paths.

To see that this maximum is achieved, suppose for contradiction that S_f is a finite initial segment of a path S in the tree and there is no σ in $\mathscr{L}_A$ such that $\mathrm{PA} + S_f + \sigma$ and $\mathrm{PA} + S_f + \neg\sigma$ are consistent. Thus we have that the theory $\mathrm{PA} + S_f$ is complete. Therefore $\mathrm{PA} + S_f$ is decidable, for given a sentence σ in $\mathscr{L}_A$ we can search the proofs of $\mathrm{PA} + S_f$ for a proof of σ or a proof of $\neg\sigma$. Therefore $\mathrm{PA} + S_f$ has recursive models. We know from Tennenbaun (Theorem A.1.6) that the only recursive model of PA is $\mathbb{N}$, so therefore $\mathrm{PA} + S_f$ has only one model $\mathbb{N}$ and S_f is a definition of truth for $\mathbb{N}$. This can be defined in $\mathscr{L}_A$, as S_f is finite, which contradicts Tarski's Theorem. □

This theorem is highly interesting, as it tells us that a satisfaction class does not provide *the* definition of truth for a model, but merely *a* definition of truth. Thus, we have answered question 2 in the negative, that satisfaction classes are not unique.

We now state and prove Lachlan's Theorem, which tells us that not all models have a satisfaction class. In particular, Lachlan's theorem tells us that a model M has a satisfaction class only if it is recursively saturated. Thus, this theorem provides an answer to question 1, and not all models have a satisfaction class. We provide a sketch of the proof here, the full details of the proof can be found in *Models of Peano Arithmetic* [12, Pages 242-246].

Theorem E.2.9 (Lachlan's)**.** *If $M \models \mathrm{PA}$ has a full satisfaction class S, then M is recursively saturated.*

Proof Sketch. Suppose that M has a full satisfaction class S and, for contradiction, that M is not recursively saturated. Then there is a recursive sequence of standard $\mathscr{L}_A$ formulas $\theta_0(\bar{x}), \theta_1(\bar{x}), \ldots, \theta_n(\bar{x}), \ldots$, where $n \in \mathbb{N}$, such that the sets $B_i = \{\bar{x} \in M : M \models \theta_i(\bar{x})\}$ form a partition of M.

We now consider the sets $C_i = \{x \in M : M \models S(\gamma_i, x)\}$ for each $i \in M$. We define γ [12, Page 244] in such a way so that we have the following informal definition of the sets C_i:

$$C_0 = \varnothing$$

$$C_{i+1} = \begin{cases} B_1 & : \text{if } C_i = \varnothing \\ B_{j+1} & : \text{if } j \text{ is the least } j \in \mathbb{N} \text{ such that } B_j \cap C_i \neq \varnothing \\ \varnothing & : \text{if no such } j \text{ exists} \end{cases}$$

This definition tells us that we have $C_i = B_i$ for all $i \in \mathbb{N}$, since all the sets B_i are disjoint from each other. Clearly $C_i \neq \varnothing$ for $i > 0$ since the sets B_i form a partition of M. We thus have that for all $i \in M$ that $C_i = B_j$ for some $j \in \mathbb{N}$. We further have that for nonstandard $i \in M$, $C_i = B_j$ implies that $C_{i-1} = B_{j-1}$ since we know j is the least natural number such that $C_{i-1} \cup B_{j-1} \neq \varnothing$ and all the sets B_n are disjoint, so we are forced by the definition that $C_{i-1} = B_{j-1}$. Now take some nonstandard $a \in M$. We know that $C_a = B_j$ for some $j \in \mathbb{N}$. However, we now get that $C_{a-j+1} = B_1$ by the above remarks which is a contradiction, since we know $a - j + 1 > 0$. Thus we cannot define the sets C_i and M has no satisfaction class which would allow us to do this [12, Pages 242-246]. □

Whilst Lachlan's theorem tells us that if a model has a satisfaction class it is recursively saturated, it does not provide the strongest condition for uncountable models. Smith's theorem does this, but to phrase this correctly we require the following definition on classes.

Definition E.2.10. *A class is a set $C \subseteq M$ such that for all $m \in M$ the set $\{c \in C : c < m\}$ is definable in M. A model is rather classless if it has no proper classes [19, Definition 2.9 and 2.14].*

We now have Smith's theorem, which says that no rather classless model has a satisfaction class. The proof of this can be found in *Nonstandard Syntax & Semantics & Satisfaction Classes in Models of Arithmetic* [19, Page 169].

Theorem E.2.11 (Smith's Theorem)**.** *If $M \vDash \mathrm{PA}$ has a full satisfaction class S, then (M, S) has an $\mathscr{L}_S$-definable proper class [19, Theorem 3.11].*

With an answer to questions 1 and 2 provided, we can now explore the answer to question 3. It turns out for countable, recursively saturated models we can find satisfaction classes. This is known as Kotlarski, Krayewski and Lachlan's (KKL's) Theorem.

E.3 M-Logic

Whilst we have proven Lachlan's theorem, to prove KKL's theorem, which is almost a converse to Lachlan's theorem, we require two new systems of logic. We define the first of these in this section, M-Logic. M-Logic is of interest as it is a logic of sentences in a specific model M and has axioms chosen that can prove the sentences that hold in M. We then have the associated interpretation of M-Logic, that of a structure satisfying the Tarski axioms.

Before exploring these axioms, it is worth noting that by expanding these axioms we are able to increase the strength of M-Logic and thus build structures which satisfy more than the Tarski axioms. In this way, we can build satisfaction classes which we satisfy certain sentences. Thus, the system presented below can be considered as a base theory of M-Logic, and we have various M-Logics which are defined from these. We explore this idea in more detail in Section 2.1.

We begin by exploring this base theory, defining its syntax and semantics, and proving it is sound and complete. For the rest of this section, we fix a nonstandard model, M, of Peano Arithmetic.

E.3.1 Syntax of M-Logic

We shall first define the mechanisms of M-Logic through its language and provability rules.

Definition E.3.1 ($^*\mathscr{L}_A(M)$)**.** *Fix a model $M \models \mathrm{PA}$. The language we work* *in for M-Logic is the language:*

$$^*\mathscr{L}_A(M) = \{x \in M : M \vDash x \text{"is the Gödel Number of a sentence of } \mathscr{L}\text{"}\}.$$

It is useful to note that we can define all numbers $m \in M$ in $^*\mathscr{L}_A(M)$. To do this we take the canonical term of these numbers, as per Definition E.1.1.

With our language in place, we now define the parameters of M-Logic.

Definition E.3.2 (Sequents). *A sequent Δ is a finite set of (Gödel Numbers of) sentences of $^*\mathscr{L}_A(M)$. If a sequent contains only one sentence σ, then rather than writing this sequent as $\{\sigma\}$, for convenience, we just write σ. If Δ is a sequent, then $\neg\Delta$ is the set of (Gödel Numbers of) the negation of the sentences in Δ [12, Definition 15.8].*

Interestingly, this means that as M-Logic is defined over sentences, we need no free variabes. With this background in place we now define provability for M-Logic.

Definition E.3.3 (Provability). *We say $\Gamma \vdash_M \Delta$, where Γ and Δ are sequents if the conjunction of all sentences of Γ implies the disjunction of all sentences of Δ by the structural and provability rules of M-Logic. We write $\varnothing \vdash_M \Delta$, or just $\vdash_M \Delta$, to mean there is a proof of the disjunction of all sentences of Δ from the rules of M-Logic. For simplicity, we write $\vdash_M \Delta \cup E$ as $\vdash_M \Delta, E$ [12, Definition 15.10].*

This also tells us that there is a symmetry between $\Gamma \vdash_M \Delta$ and $\vdash_M \Delta, \neg\Gamma$.

Lemma E.3.4. $\Gamma \vdash_M \Delta$ *if and only if* $\vdash_M \Delta, \neg\Gamma$.

Proof. Suppose $\Gamma \vdash_M \Delta$. Thus, either $\vdash_M \Delta$ or some sentence in Γ is not implied by M-Logic, i.e. $\vdash_M \neg\Gamma$. Therefore, $\vdash_M \Delta, \neg\Gamma$. □

We now need to set up these rules of M-Logic. We set them up in two sections, the first of these are the following structural rules which tell us that the sequents behave like sets. The first says we can add identical sentences, the second says we can remove identical sentences and the third says we can rearrange sentences at will.

Definition E.3.5 (Structural rules). *For any sequent* Δ *and any sentences* σ *and* τ *in* ${}^*\mathscr{L}_A(M)$*:*

1. *If* $\vdash_M \Delta, \sigma$*, then* $\vdash_M \Delta, \sigma, \sigma$
2. *If* $\vdash_M \Delta, \sigma, \sigma$*, then* $\vdash_M \Delta, \sigma$
3. *If* $\vdash_M \Delta, \sigma, \tau$*, then* $\vdash_M, \Delta, \tau, \sigma$ *[12, Page 248]*

We now define the provability rules of M-Logic.

Definition E.3.6 (Provability rules). *Let* Δ, E *be sequents,* σ *and* τ *be* ${}^*\mathscr{L}_A(M)$*-sentences and* $\theta(x)$ *be a* ${}^\star\mathscr{L}_A$*-formula.The provability rules of M-Logic are:*

4. *We have that* $\vdash_M \sigma, \neg\sigma$
5. *If* t *and* s *are terms, then:*
   (a) *If* $M \models val(t) = val(s)$*, then* $\vdash_M t = s$
   (b) *If* $M \models val(t) \neq val(s)$*, then* $\vdash_M t \neq s$

(c) *If* $M \vDash val(t) < val(s)$, *then* $\vdash_M t < s$

(d) *If* $M \vDash val(t) \geqslant val(s)$, *then* $\vdash_M \neg(t < s)$

6. *(Weakening) If* $\vdash_M \Delta$, *then* $\vdash_M \Delta, \sigma$

7. *(Cut) If* $\vdash_M \Delta, E$ *and* $\vdash_M \Delta, \neg E$, *then* $\vdash_M \Delta$

8. *If* $\vdash_M \Delta$, *then* $\vdash_M \neg\neg\Delta$

9. (a) *If* $\vdash_M \Delta, \sigma$ *or* $\vdash_M \Delta, \tau$, *then* $\vdash_M \Delta, \sigma \vee \tau$

   (b) *If* $\vdash_M \Delta, \neg\sigma$ *and* $\vdash_M \Delta, \neg\tau$, *then* $\vdash_M \Delta, \neg(\sigma \vee \tau)$

10. (a) *If* $\vdash_M \Delta, \sigma$ *and* $\vdash_M \Delta, \tau$, *then* $\vdash_M \Delta, \sigma \wedge \tau$

    (b) *If* $\vdash_M \Delta, \neg\sigma$ *or* $\vdash_M \Delta\neg\tau$, *then* $\vdash_M \Delta, \neg(\sigma \wedge \tau)$

11. *For all terms* t, s, r :

    (a) *(Reflexivity of =)* $\vdash_M t = t$

    (b) *(LEM for =)* $\vdash_M t = s, \neg(t = s)$

    (c) *(Transitivity of =)* $\vdash_M \neg(t = s), \neg(s = r), t = r$

12. *Let* $t_1, t_2, \ldots, t_n$ *and* $s_1, s_2, \ldots, s_n$ *be terms, then:*

$$\vdash_M \neg t_1 = s_1, \neg t_2 = s_2, \ldots, \neg t_n = s_n, \neg\theta(t_1, t_2, \ldots, t_n), \theta(s_1, s_2, \ldots, s_n)$$

13. *(M-Rule a)*

    (a) *If* $\vdash_M \Delta, \theta(t)$, *then* $\vdash_M \Delta, \exists x\theta(x)$

    (b) *If* $\vdash_M \Delta, \neg\theta(t)$ *for all terms* $t \in M$, *then* $\vdash_M \Delta, \neg\exists x\theta(x)$

14. (M-Rule b)

(a) If $\vdash_M \Delta, \theta(t)$ *for all* $t \in M$*, then* $\vdash_M \Delta, \forall x\theta(x)$

(b) If $\vdash_M \Delta, \neg\theta(t)$*, then* $\vdash_M \Delta, \neg\forall x\theta(x)$ *[12, Definition 15.10]*

We see hence that proofs in M-Logic correspond to trees that are infinite in both width and depth. We have proofs of infinite width from the M-Rules and this entails proofs of infinite depth. If in a subtree of infinite width there is an ordering of the paths whereby their lengths are strictly increasing, then the whole subtree will have infinite depth.

E.3.2 Semantics of M-Logic

With the rules of M-Logic in place we are able to define the semantics associated with M-Logic and what it means for M-Logic to be consistent, sound and complete. In particular, the interpretation of M-Logic is that it is consistent only if there is a full satisfaction class over M.

Definition E.3.7 (Semantics)**.** *We have that* $\vDash_M \Delta$*, the disjunction of* Δ *is semantically valid in* M*-Logic, if every full satisfaction class over* M *makes some sentence* $\delta \in \Delta$ *true [12, Definition 15.9].*

Definition E.3.8 (Consistent)**.** *We say that* M*-Logic is semantically consistent if there is some full satisfaction class* S *over* M*. We say* M*-Logic is syntactically consistent if* $\nvdash_M \varnothing$*, or sometimes just as* $\nvdash_M$*. This is because if* $\vdash_M \varnothing$*, then by weakening we have that* M*-Logic can prove any sentence [12, Definition 15.13].*

E.3.3 M-Logic is Sound and Complete

We can prove that M-Logic is sound and complete, we begin by defining what it means for M-Logic to be sound, and then proving this.

Definition E.3.9 (Sound)**.** *M-Logic is sound if whenever $\vdash_M \Delta$ and S is a full satisfaction class over M, then S makes at least one $\delta \in \Delta$ true.*

Theorem E.3.10. *For every model M, M-Logic is sound [12, Proposition 15.11].*

Proof. We perform induction on the length of derivation of a proof of M-Logic. We hence consider each rule in turn and suppose that it was the last step in the derivation.

Rule 1. If our last step was to derive $\vdash_M \Delta, \sigma, \sigma$, from $\vdash_M \Delta, \sigma$, then by our inductive hypothesis we know that every satisfaction class over M makes either σ or some sentence $\delta \in \Delta$ true. Thus every satisfaction class over M makes some sentence $\delta \in \Delta$ true, σ true or σ true by triviality.

Rule 2. If our last step was that $\vdash_M \Delta, \sigma, \sigma$ implies $\vdash_M \Delta, \sigma$, then by inductive hypothesis we have that every satisfaction class over M makes some $\delta \in \Delta$ true, σ true or σ true. Thus, every satisfaction class over M makes some $\delta \in \Delta$ true or σ true.

Rule 3. If we derived $\vdash_M \Delta, \tau, \sigma$ from $\vdash_M, \Delta, \sigma, \tau$. then by induction every satisfaction class over M makes some $\delta \in \Delta$ true, σ true or τ true and thus we are done.

Rule 4. If our last step was $\vdash_M \Delta, \sigma, \neg\sigma$ from $\vdash_M \Delta$, then by induction every satisfaction class over M makes some $\delta \in \Delta$ true, and by definition a full satisfaction class makes σ or $\neg\sigma$ true, so every satisfaction class makes some $\delta \in \Delta \cup \{\sigma, \neg\sigma\}$ true.

Rule 5. If we last derived $\vdash_M t = s$, then we know $M \vDash val(t) = val(s)$. Thus, since a satisfaction class satisfies the Tarski Axiom 1, it makes the sentence $t = s$ true. It is a simple exercise to show this holds for the other cases of Rule 5 as well.

Rule 6. If we last derived $\vdash_M \Delta, \sigma$ from $\vdash_M \Delta$, then we know by induction that every satisfaction class over M makes some $\delta \in \Delta$ true, so every satisfaction class trivially makes some $\delta \in \Delta$ or σ true.

Rule 7. If our last step was to derive $\vdash_M \Delta$ from $\vdash_M \Delta, E$ and $\vdash_M \Delta, \neg E$, then we know every saitsfaction class makes some $\delta \in \Delta \cup E$ true and every satisfaction class makes some $\delta \in \Delta \cup \neg E$ true. We know that no satisfaction class can make both σ and $\neg\sigma$ true by Tarski Condition 5, so every satisfaction class makes some $\delta \in \Delta$ true.

Rule 8. If we last derived $\vdash_M \neg\neg\Delta$ from $\vdash_M \Delta$ then we know via induction that every satisfaction class makes some $\delta \in \Delta$ true. We thus have that every satisfaction class makes $\neg\neg\delta$ true for some $\delta \in \Delta$, by applying the Tarski Condition 5 twice. Therefore every satisfaction class makes some $\delta \in \neg\neg\Delta$ true.

Rule 9. (a) If we last derived $\vdash_M \Delta, \sigma \vee \tau$ from $\vdash_M \Delta, \sigma$ or $\vdash_M \Delta, \tau$, then we know every satisfaction class over M makes some $\delta \in \Delta \cup \{\sigma\}$

true or makes some $\delta \in \cup\{\tau\}$ true. Thus every satisfaction class over M makes some $\delta \in \Delta$ true or σ or τ true. Thus, we have that M makes some $\delta \in \Delta \cup \{\sigma \vee \tau\}$ true by Tarski Condition 4.

(b) If our last step was to derive $\vdash_M \Delta, \neg(\sigma \vee \tau)$ from $\vdash_M \Delta, \neg\sigma$ and $\vdash_M \Delta, \neg\tau$, then by induction we know that every satisfaction class makes some $\delta \in \Delta \cup \{\neg\sigma\}$ true and some $\delta \in \Delta \cup \{\neg\tau\}$ true. If every satisfaction class over M makes $\sigma \vee \tau$ true, then by Tarski Condition 4 every satisfaction class over M makes σ true or τ true, so we hence have that every satisfaction class over M makes some $\delta \in \Delta$ true. Therefore, by Tarski Condition 5, we have that every satisfaction class makes some $\delta \in \Delta \cup \{\neg(\sigma \vee \tau)\}$ true.

Rule 10. (a) If we last derived $\vdash_M \Delta, \sigma \wedge \tau$ from $\vdash_M \Delta, \sigma$ and $\vdash_M \Delta, \tau$, then we know every satisfaction class over M makes some $\delta \in \Delta \cup \{\sigma\}$ true and makes some $\delta \in \cup\{\tau\}$ true. Thus every satisfaction class over M makes some $\delta \in \Delta$ true or, σ true and τ true. Thus, we have that M makes some $\delta \in \Delta \cup \{\sigma \wedge \tau\}$ true by Tarski Condition 3.

(b) If our last step was to derive $\vdash_M \Delta, \neg(\sigma \wedge \tau)$ from $\vdash_M \Delta, \neg\sigma$ or $\vdash_M \Delta, \neg\tau$, then by induction we know that every satisfaction class makes some $\delta \in \Delta \cup \{\neg\sigma\}$ true or some $\delta \in \Delta \cup \{\neg\tau\}$ true. If every satisfaction class over M makes $\sigma \wedge \tau$ true, then by Tarski Condition 3 every satisfaction class over M makes

σ true and τ true, so we hence have that every satisfaction class over M makes some $\delta \in \Delta$ true. Therefore, by Tarski Condition 5, we have that every satisfaction class makes some $\delta \in \Delta \cup \{\neg(\sigma \wedge \tau)\}$ true.

Rule 11. (a) If we have derived $\vdash_M \Delta, t = t$ from $\vdash_M \Delta$, then by Tarski Condition 1 since $val(t) = val(t)$ we have that every satisfaction class over M makes some $\delta \in \Delta, t = t$ true.

(b) If we last derived $\vdash_M, \Delta \neg t = s, s = t$ from $\vdash_M \Delta$, then since $val(t) = val(s)$ or $\neg(val(t) = val(s))$ we have by Tarski conditions 1 and 5 that every satisfaction class over M makes some $\delta \in \Delta, t = s, \neg(s = t)$ true.

(c) Suppose that we last derived $\vdash_M \Delta, neg(t = s), \neg(s = r), t = r$ from $\vdash_M \Delta$. We know that $t \neq s$, $s \neq r$ or $t = r$, since if $t = s$ and $s = r$, then $t = r$ by transitivity. Thus, by Tarski Condition 1) and our inductive hypothesis we know that every satisfaction class over M makes some $\delta \in \Delta \cup \{\neg(t = s), \neg(s = r), t = r\}$ true.

Rule 12. Suppose that our last step was to derive $\vdash_M \Delta, \neg t_1 = s_1, \neg t_2 = s_2, \ldots, \neg t_n = s_n, \neg\theta(t_1, t_2, \ldots, t_n), \theta(s_1, s_2, \ldots, s_n)$ from $\vdash_M \Delta$. We know that if $t_i = s_i$ for all i and $\theta(t_1, t_2, \ldots, t_n)$, then $\theta(s_1, s_2, \ldots, s_n)$, so by Tarski Conditions 1) and 5) and our inductive hypothesis we have that every full satisfaction class over M makes some $\delta \in \Delta \cup \{\neg(t_1 = s_1), \ldots, \neg(t_n = s_n), \theta(t_1, \ldots, t_n), \neg\theta(s_1, \ldots, s_n)$.

Rule 13. (a) If we last derived $\vdash_M \Delta, \exists x\theta(x)$ from $\vdash_M \Delta, \theta(t)$, then we know that every full satisfaction class over M makes some $\delta \in \Delta$ true or makes $\theta(t)$ true for some term t. Therefore, every full satisfaction class over M makes some $\delta \in \Delta$ true or makes $\exists x\theta(x)$ true by Tarski Condition 6.

(b) If our last step was $\vdash_M \Delta, \neg\exists x\theta(x)$ from $\vdash_M \Delta, \neg\theta(t)$ for all terms t. Then by our inductive assumption and Tarski Conditions 5) and 6) we have that every full satisfaction class over M makes some $\delta \in \Delta \cup \{\neg\exists x\theta(x)\}$ true.

Rule 14. (a) If we last derived $\vdash_M \Delta, \forall x\theta(x)$ from $\vdash_M \Delta, \theta(t)$ for all t, then we know that every full satisfaction class over M makes some $\delta \in \Delta$ true or makes $\theta(t)$ true for all terms t. Therefore, every full satisfaction class over M makes some $\delta \in \Delta$ true or makes $\forall x\theta(x)$ true by Tarski Condition 7.

(b) If our last step was $\vdash_M \Delta, \neg\forall x\theta(x)$ from $\vdash_M \Delta, \neg\theta(t)$. then by our inductive assumption and Tarski Conditions 5) and 7) we have that every full satisfaction class over M makes some $\delta \in \Delta \cup \{\neg\forall x\theta(x)\}$ true.

□

We now define what it means for M-Logic to be complete and prove this.

Definition E.3.11 (Complete). *M-Logic is complete if every full satisfaction class makes some δ true, then $\vdash_M \delta$.*

Theorem E.3.12. *For every countable model M, M-Logic is complete.*

Proof. We prove this by contrapositive. Suppose that $\Gamma \nvdash_M \varnothing$. We aim to show that this entails $\Gamma \nvDash_M \varnothing$,

If $\Gamma \nvdash_M \varnothing$, then for all $\mathscr{L}_A$-sentences σ, it is not the case that both σ and $\neg\sigma$ are in Γ, since $\Gamma \vdash_M \sigma, \neg\sigma$ by Rule 4) and if this were the case, then by applying the cut rule, Rule 8) and the cut rule again we have that $\Gamma \vdash_M \varnothing$, which is not the case.

Given Γ we can extend Γ to contain either any ${}^\star\mathscr{L}_A$-sentence σ or $\neg\sigma$. This is because if $\Gamma, \sigma \vdash_M \varnothing$, then $\Gamma \vdash_M \neg\sigma$, so we can consistently add $\neg\sigma$ to Γ, if we cannot add σ. Similarly, given an ${}^\star\mathscr{L}_A$-formula $\phi(x)$, we can either add $\phi(a)$ for some constant $a \in M$ or add $\neg\phi(a)$ for all constants a which entails that we can add $\neg\exists x\phi(x)$ to Γ. This is since if $\Gamma, \phi(a) \vdash_M \varnothing$, for all constants a, then $\Gamma \vdash_M \neg\phi(a)$ for all a, so we can add $\neg\phi(a)$ for all constants a. This then tells us, by the M-Rule a), $\Gamma \vdash_M \neg\exists x\phi(x)$, so we can consistently add this to Γ.

We thus enumerate all ${}^\star\mathscr{L}_A$-sentences and formulas, using the countability of M, and inductively construct a $\Sigma \supseteq \Gamma$, by letting $\Sigma_0 = \Gamma$ and dovetailing the following processes together:

- Given Σ_i take $a \in M$ and a formula $\phi(x)$. If $\Sigma_i, \phi(a) \nvdash_M \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\phi(a)\}$, else take $\Sigma_{i+1} = \Sigma_i$.

- Given Σ_i take $a \in M$ and a formula $\phi(x)$. If $\Sigma_i, \neg\phi(a) \nvdash_M \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\neg\phi(a)\}$, else take $\Sigma_{i+1} = \Sigma_i$.

- Given Σ_i take a sentence σ. If $\Sigma_i, \sigma \nvdash_M \varnothing$, then we let $\Sigma_{i+1} = \Sigma_i \cup \{\sigma\}$, else take $\Sigma_{i+1} = \Sigma_i \cup \{\neg\sigma\}$.

We have that our final set $\Sigma = \bigcup_{i\in\mathbb{N}} \Sigma_i$ satisfies the following four properties:

1. All finite sets $\Delta \subseteq \Sigma$ have the property that $\Sigma_n \nvdash_M \varnothing$.

2. For all ${}^\star\mathscr{L}_A$-sentences σ either $\sigma \in \Sigma$ or $\neg\sigma \in \Sigma$.

3. If for a ${}^\star\mathscr{L}_A$-formula $\phi(x)$ we have that $\phi(a) \in \Sigma$ for all $a \in M$, then $\Sigma_n \vdash_M \forall x\phi(x)$ for some finite $\Sigma_n \subseteq \Sigma$.

4. If for a ${}^\star\mathscr{L}_A$-formula $\phi(x)$ we have that $\neg\phi(a) \in \Sigma$ for all $a \in M$, then $\Sigma_n \vdash_M \neg\exists x\phi(x)$ for some finite $\Sigma_n \subseteq \Sigma$.

Σ satisfies property 1 because if $\Delta \subseteq \Sigma$ is finite, then $\Delta \subseteq \Sigma_n$ for some n and thus, since $\Sigma_n \nvdash_M \varnothing$ we have by repeated use of the weakening rule that $\Delta \nvdash_M \varnothing$.

Σ satisfies property 2, since each ${}^\star\mathscr{L}_A$-sentence σ is σ_i for some i, and so by our inductive construction Σ contains either σ or $\neg\sigma$.

Σ satisfies property 3, since if $\phi(a) \in \Sigma$ for all $a \in M$, then at some stage in our construction it is inconsistent to add $\neg\phi(a)$ to Σ_n for all $a \in M$. Thus, $\Sigma_n \vdash_M \phi(a)$ for all $a \in M$ and by the M-Rule b $\Sigma_n \vdash_M \forall x\phi(x)$.

Similarly, Σ satisfies property 4, since if $\neg\phi(a) \in \Sigma$ for all $a \in M$, then at some stage in our construction it is inconsistent to add $\phi(a)$ to Σ_n for all $a \in M$. Thus, $\Sigma_n \vdash_M \neg\phi(a)$ for all $a \in M$ and by the M-Rule a $\Sigma_n \vdash_M \neg\exists x\phi(x)$.

We now prove that Σ gives a full satisfaction class S over M. Clearly, if some finite $\Delta \subseteq \Sigma$ is such that $\Delta \vdash_M \sigma$, then $\sigma \in \Sigma$. This, combined with the M-Rules, proves that Σ satisfies Tarski Condition 1) and 2). We also know from this that Σ satisfies Tarski Condition 3) and 4) by provability rules 9) and 10). We have that Σ satisfies Tarski Condition 5) by our property 2. Finally, Σ satisfies Tarski Conditions 6) and 7) via properties 3 and 4.

Thus there is a full satisfaction class Σ over M which makes each $\gamma \in \Gamma$ false, and M-Logic is complete. [12, Proposition 15.14] $\square$

Thus we have that if M-Logic is consistent for a countable model M, then M-Logic has a model and this model corresponds to a full satisfaction class over M. However, to prove that M-Logic is consistent we work in an adaptation of M-Logic, called FA-Logic.

E.4 FA-Logic

FA-Logic (Finite Approximation Logic) is an adaptation of M-Logic, that includes formulas and quantifiers. Thus, FA-Logic is defined over ${}^*\mathscr{L}_A(M)$ formulas, rather than sentences. We introduce FA-Logic because it can approximate M-Logic, and thus allows us to prove the consistency of M-Logic, and hence KKL's Theorem that some models have a full satisfaction class. We start off by defining the syntax for FA-Logic.

E.4.1 Provability in FA-Logic

The provability rules for FA-Logic are the same as those for M-Logic, with the exception that the M-Rules are replaced by rules involving formulas.

Definition E.4.1. *The rules of FA-Logic are the structural and provability rules of M-Logic, without both M-Rules. They are replaced by:*

13. *(a) If for a term t, if $\vdash_{\mathrm{FA}} \Delta, \theta(t)$, then $\vdash_{\mathrm{FA}} \Delta, \exists x \theta(x)$.*

(b) *If for a variable* v, $\vdash_{\mathrm{FA}} \Delta, \neg\theta(v)$, *then* $\vdash_{\mathrm{FA}} \Delta, \neg\exists x\theta(x)$.

14. (a) *If for a variable* v, $\vdash_{\mathrm{FA}} \Delta, \theta(v)$, *then* $\vdash_{\mathrm{FA}} \Delta, \forall x\theta(x)$.

    (b) *If for a term* t, $\vdash_{\mathrm{FA}} \Delta\neg\theta(t)$, *then* $\vdash_{\mathrm{FA}} \Delta, \neg\forall x\theta(x)$ *[12, Definition 15.15].*

Definition E.4.2 (Provability). *For any sequent of formulas from* ${}^*\mathscr{L}_A(M)$ *we say that* $\vdash_{\mathrm{FA}} \Delta$ *if there is a proof of the disjunction of* Δ *by the rules of FA-Logic. We use the same conventions as* M*-Logic are say* $\Delta \vdash_{\mathrm{FA}} E$ *means that* $\vdash_{\mathrm{FA}} E, \neg\Delta$ *(the conjunction of* Δ *implies the disjunction of* E*).*

Thus, whereas a proof in M-Logic can be an infinite tree, we have that all proofs in FA-Logic are finite trees of ${}^*\mathscr{L}_A(M)$-formulas.

With the syntax of FA-Logic in place we are now able to explore its use in proving the consistency of M-Logic for certain models, which is KKL's Theorem.

E.4.2 KKL's Theorem

Kotlarski, Krayewski and Lachlan's Theorem is a major result in the theory of satisfaction classes and says that every countable, recursively saturated nonstandard model has a satisfaction class. We shall outline the proof of this below. To start with, we define the set of true sentences of a model M below as $\mathrm{Diag}(M)$.

Definition E.4.3. *The standard full diagram of a model* M *is denoted by* $\mathrm{Diag}(M)$ *and is* $\{\theta(\bar{a}) : M \vDash \theta(\bar{a})$ *and* $\theta(\bar{x})$ *is an* $\mathscr{L}_A(M)$*-formula*$\}$ *[12, Definition 15.16].*

From this definition we can think of Diag(M) as the set of all true sentences of our model $M \models$ PA. Hence we know that this is consistent and we can use this to derive that FA-Logic is consistent.

Theorem E.4.4. *If* $M \neq \mathbb{N}$*, then* Diag(M) $\nvdash_{\text{FA}} 0 = 1$*.*

Proof Sketch. If this were not the case, then there would be finite proof of ${}^*\mathscr{L}_A(M)$-formulas showing that $0 = 1$. Thus, the proof is bounded above in complexity by Σ_n for some $n \in \mathbb{N}$. We can then use a weakened Tarski truth predicate Tr_{Σ_n} which is conservative in PA to show that each of the sentences and deductions in the proof are true in M for this truth predicate. However, we have $\neg Tr_{\Sigma_n}(0 = 1, a)$ which is a contradiction. [12, Page 252] □

We now demonstrate that FA-Logic can simulate M-Logic. In particular that it is able to simulate the M-Rules 13 and 14. We can then combine these two theorems to show that M-Logic is thus consistent.

Theorem E.4.5. *If* $M \neq \mathbb{N}$ *and* M *is recursively saturated, then for any sequent* Δ *and* ${}^*\mathscr{L}_A(M)$*-formula* θ *we have that:*

a) If Diag(M) $\vdash_{\text{FA}} \Delta, \neg\theta(a)$ *for all* $a \in M$*, then* Diag(M) $\vdash_{\text{FA}} \Delta, \neg\exists x\theta(x)$*.*

b) If Diag(M) $\vdash_{\text{FA}} \Delta, \theta(a)$ *for all* $a \in M$*, then* Diag(M) $\vdash_{\text{FA}} \Delta, \forall x\theta(x)$ *[12, Lemma 15.18]*

Proof Sketch. Here we shall provide a sketch of case a). The proof for the case b) is similar and the full details for the proof of this case can be found in Chapter 15 of *Models of Peano Arithmetic* [12, Pages 253-256].

The idea of this proof is that since proofs in FA-Logic are finite we can verify if a proof in FA-Logic is a valid proof and write, in $\mathscr{L}_A$, $\vdash^p_{\text{FA}} \Gamma$ to say that p is a correct proof in FA-Logic of Γ. Since there are countably many such proofs we can recursively enumerate them and translate $\text{Diag}(M) \vdash_{\text{FA}} \Delta, \neg\theta(a)$ for all $a \in M$ into the language $\mathscr{L}^{rec}_{\omega_1\omega}$ as:

$$\forall a \in M \bigvee_p \text{Diag}(M) \vdash^p_{\text{FA}} \Delta, \neg\phi(a).$$

We can then use recursive saturation to get that there are hence finitely many proofs $p_1, \ldots, p_n$ such that:

$$\forall a \in M \bigvee_{p_i \in \{p_1,\ldots,p_n\}} \text{Diag}(M) \vdash^{p_i}_{\text{FA}} \Delta, \neg\phi(a).$$

Now, we can use the rules of FA-Logic to show that $\bigvee_{p_i \in \{p_1,\ldots,p_n\}} \text{Diag}(M) \vdash^{p_i}_{\text{FA}} \Delta, \neg\exists x\phi(x)$ and working backwards we prove that $\text{Diag}(M) \vdash_{\text{FA}} \Delta, \neg\exists x\theta(x)$.

This appeal to recursive saturation is not quite right, since we are ranging over infinitely many elements of M. To get around this issue we use the notion of a proof template as defined in *Models of Peano Arithmetic* [12, Definition 15.19]. Informally, a proof template is a (standard) finite object which is the structure of an FA-Logic proof with gaps for elements of M to be inserted into. We then no longer range over all elements of M, since the proof templates contain no reference to these numbers. □

These two theorems allow us to prove KKL's theorem, that any countable recursively saturated nonstandard model has a full satisfaction class.

Theorem E.4.6 (KKL's Theorem)**.** *Let* $M \vDash \text{PA}$ *be countable, nonstan-*

dard and recursively saturated. Then M has a full satisfaction class S [12, Theorem 15.21].

Proof. If M-Logic were inconsistent, then $\vdash_M 0 = 1$, and thus, by Theorem E.4.5 since FA-Logic can mimic M-Logic, we have that $\mathrm{Diag}(M) \vdash_{\mathrm{FA}} 0 = 1$. However, by Theorem E.4.4 we know that this is not the case. Therefore M-Logic is consistent and since M is countable we know that hence by completeness M has a full satisfaction class S. $\square$

This surprising theorem tells us that we can find a definition of truth for nearly any countable model of PA. This is because we know that any model $m \vDash$ PA has an elementary end extension which is recursively saturated, i.e. there exists $N \succ_e M$ where N is recursively saturated. This is highly useful, but is not as strong a result as it could be.

E.5 Strengthening Satisfaction

We have seen previously that every countable recursively saturated nonstandard model $M \vDash$ PA has a full satisfaction class (Theorem E.2.9), however we have also seen that if a countable model has a satisfaction class, then it has $2^{\aleph_0}$ such satisfaction classes (Theorem E.2.8). In other words, we have many different definitions of truth for these models, which disagree on the truth of various nonstandard sentences. This leads to a rich and interesting theory, but means that we have many satisfaction classes for a model M which set some intuitively true sentence as false, or, some intuitively false sentence as true.

E.5.1 Pathologies

There are various nonstandard sentences which we believe we know the truth of. Examples are easy to think of, a canonical one being the sentence:

$$1 = 1 \wedge (1 = 1 \wedge (1 = 1 \wedge (1 = 1 \wedge (\cdots \wedge (1 = 1) \cdots))))$$

which has a nonstandard number of disjuncts. We will write this sentence as $\bigwedge_a 1 = 1$, where a is the nonstandard number of disjuncts. This sentence is intuitively true, as $1 = 1$, no matter how many times we examine it. The similarly notated sentence $\bigvee_a 0 = 1$ is also intuitively false. One facet of satisfaction classes as a definition of truth is that they can interpret the sentence above as true, and various intuitively true sentences, such as $\bigwedge_a 1 = 1$ as false. We shall examine some examples of this below in more detail.

Theorem E.5.1. *Let $M \models \mathrm{PA}$ be countable, recursively saturated and $M \neq \mathbb{N}$. Then M has a satisfaction class S such that for some nonstandard $t \in M$: $(M, S) \models S(\ulcorner \bigvee_t 0 = 1 \urcorner, a)$.*

Proof Sketch. Let $M \models \mathrm{PA}$ be as above (countable, nonstandard and recursively saturated). To prove that there is a satisfaction class S over M making $\bigvee_t 0 = 1$ true we show that the sentence is M-consistent. This is sufficient by the soundness and consistency theorems. We have that $\bigvee_t 0 = 1 \vdash_M \varnothing$ if and only if $\vdash_M \neg \bigvee_t 0 = 1$ by Lemma E.3.4. However, proving this sentence in M-Logic requires us to show $\vdash_M \neg(0 = 1)$ and $\vdash_M \neg \bigvee_{t-1} 0 = 1$. Proving the latter half of this, again requires $\vdash_M \neg \bigvee_{t-2} 0 = 1$. At no stage are we able to prove the latter half of this nonstandard disjunction. Therefore, since we cannot prove $\vdash_M \neg \bigvee_t 0 = 1$ we also cannot prove $\bigvee_t 0 = 1 \vdash_M \varnothing$. $\square$

Satisfaction classes can also take sentences which are intuitively true and in a similar manner believe that these sentences are false. We consider an example of this below.

Theorem E.5.2. *Let $M \models \mathrm{PA}$ be countable, recursively saturated and $M \neq \mathbb{N}$. Then M has a satisfaction class S such that for some nonstandard $t \in M$: $(M, S) \not\models S(\ulcorner \bigwedge_t 1 = 1 \urcorner, a)$.*

Proof Sketch. We prove this similarly to the above. We know that $(M, S) \not\models S(\ulcorner \bigwedge_t 1 = 1 \urcorner, a)$ if and only if $(M, S) \models \neg S(\ulcorner \bigwedge_t 1 = 1 \urcorner, a)$, which by Tarski Condition 5 holds if and only if $(M, S) \models S(\ulcorner \neg \bigwedge_t 1 = 1 \urcorner, a)$. We now show that, given some set Σ which is consistent in M-Logic, $\Sigma \cup \{\bigwedge_t 1 = 1\}$ is inconsistent in M-Logic. This entails that $\bigwedge_t 1 = 1$ is not always M-consistent so there are satisfaction classes which don't contain this sentence. Using a similar argument as in the proof to Theorem E.5.1 we have that $\{\neg \bigwedge_t 1 = 1\}$ is consistent in M-Logic, but adding the sentence $\bigwedge_t 1 = 1$ to this set is clearly inconsistent in M-Logic via the cut rule. $\square$

Again, whilst a satisfaction class can believe this sentence is false, it is intuitively true, as $1 = 1$ should always hold. Thus one of the key areas of study for satisfaction classes are ways of strengthening them to reduce pathologies such as these.

There are numerous methods of strengthening satisfaction classes, including adding axioms to the language including satisfaction classes, adding further Tarski conditions and specifying that certain sentences are true in a satisfaction class.

E.5.2 Adding Axioms to $\mathscr{L}_S$

In this section we look at the option of adding further axioms to satisfaction classes. This area has been highly discussed in the literature and has produced some important theorems which we shall discuss here.

The first theory that we shall consider is that of PA(S) which is that of PA where S is a satisfaction class and all the induction axioms hold for S.

Definition E.5.3 (PA(S))**.** *We denote by* PA(S) *the axiom schema* PA + *S is a full satisfaction class for the model considered + the axiom schema of induction in the language $\mathscr{L}_S$ [12, Exercise 15.9].*

This theory, PA(S) is very strong and is able to remove all the pathologous examples, such as those considered above. We prove an example of this below.

Example E.5.4. $\mathrm{PA}(S) \vdash \neg S(\ulcorner \bigvee_t 0 = 1 \urcorner, a)$

Proof. Let $(M, S) \vDash \mathrm{PA}(S)$. We prove that for all $t \in M$ it is the case that $(M, S) \vDash \neg S(\ulcorner \bigvee_t 0 = 1 \urcorner, a)$. Clearly $(M, S) \vDash \neg S(0 = 1)$. If $(M, S) \vDash \neg S(\ulcorner \bigvee_c 0 = 1 \urcorner, a)$, then $(M, S) \vDash \neg S(\ulcorner \bigvee_{c+1} 0 = 1 \urcorner, a)$ by Tarski Condition 4. Thus, by induction in $\mathscr{L}_S$ we are done. □

Whilst this does remove pathological examples, the problem with this a base theory is that it is very strong and rules out various interesting models that we would like to consider. As an example of its strength, we prove below that it entails the consistency of PA.

Theorem E.5.5. *It is the case that* $\mathrm{PA}(S) \vdash con(\mathrm{PA})$.

Proof Sketch. We let $(M,S) \models \text{PA}(S)$ and then suppose for contradiction that $(M,S) \models \neg con(\text{PA})$. Using induction on x in $\mathscr{L}_S$ we can prove the formula:

$$\forall p{<}x \forall \alpha{<}x[\text{“}\llcorner p \lrcorner \text{ proves } \llcorner \alpha \lrcorner \text{ in PA”} \rightarrow \forall a S(\theta, a)].$$

Then, we have that $(M,S) \models \exists p[\text{“}\llcorner p \lrcorner \text{ proves } \neg(0=1) \text{ in PA”}]$, so therefore $(M,S) \models S(\ulcorner 0 = 1 \urcorner, a)$ and $(M,S) \models S(\ulcorner \neg(0=1) \urcorner, a)$ which is a contradiction. [12, Page 247] □

Given the result above, that PA(S) proves the consistency of PA, we thus would like to weaken the theory whilst still removing the pathological examples we have looked at. One attempt is to restrict the induction in $\mathscr{L}_S$ only to certain classes of formulas, given that full induction is not needed to remove these pathological examples. It turns out, however, that this theory is still highly strong. We now look at the theory $\Delta_0 - \text{PA}(S)$ which has induction only for Δ_0 formulas in $\mathscr{L}_S$.

Definition E.5.6 ($\Delta_0{-}\text{PA}(S)$)**.** *We denote by* $\Delta_0{-}\text{PA}(S)$ *the axiom schema* PA*, that* S *is a full satisfaction class for the model considered and the axiom schema of induction for all* Δ_0 *formulas of* $\mathscr{L}_S$*.*

Whilst this theory is weaker than PA(S) it is still very strong and also entails the consistency of PA as stated by Ciéśliński [4, Page 8].

Theorem E.5.7. *It is the case that* $\Delta_0 - \text{PA}(S) \vdash con(\text{PA})$*. [4, Page 8]*

The theory $\Delta_0{-}\text{PA}(S)$ is one which is obtained quite naturally in an effort to remove pathologies from satisfaction classes. When considering the pathological examples above, one notices that $0 = 1 \vee (0 = 1 \vee (\cdots \vee (0 = 1)))$

is semantically equivalent to the sentence $0 = 1$. Similarly, the sentence $1 = 1 \wedge (1 = 1 \wedge (\cdots \wedge (1 = 1)))$ is semantically equivalent to the sentence $1 = 1$. We know that a satisfaction class can deal with the truth of these standard sentences. Therefore, if a satisfaction class recognises this equivalence, then these pathological examples would be removed. Ciésliński [4] closes satisfaction classes under proof in propositional logic to remove pathological examples like this. We call this theory $\mathrm{PA} + PL(S)$.

Definition E.5.8 ($\mathrm{PA} + PL(S)$)**.** *The theory* $\mathrm{PA} + PL(S)$ *is* $\mathrm{PA} + S$ *is a full satisfaction class* $+\forall\alpha[$*"*$\llcorner\alpha\lrcorner$ *is provable in propositional logic from satisfied sentences"* $\rightarrow \forall a S(\alpha, a)]$ *[4, Page 9]*

This theory is strong enough to remove the pathological examples we have hence considered above. We prove an example of this below.

Lemma E.5.9. *For all* $a, t \in M$*,* $\mathrm{PA} + PL(S) \vdash \neg S(\ulcorner \bigvee_t 0 = 1 \urcorner, a)$

Proof. Let $(M, S) \models \mathrm{PA} + PL(S)$. We know that $(M, S) \models S(\neg(0 = 1), a)$. We can prove in propositional logic that $\neg(0 = 1) \leftrightarrow \bigvee_t \neg(0 = 1)$, therefore by Tarski Condition 5 we have that $(M, S) \models \neg S(\ulcorner \bigvee_t 0 = 1 \urcorner, a)$. □

Thus we have another successful method of removing at least some pathological examples from the theory of satisfaction classes. However, this theory is also highly strong and is actually equivalent to the theory of $\Delta_0 - \mathrm{PA}(S)$.

Theorem E.5.10. *The theory* $\mathrm{PA} + PL(S)$ *is equivalent to the theory* $\Delta_0 - \mathrm{PA}(S)$

Proof Idea. In the theory $\mathrm{PA}+PL(S)$ given any Δ_0 formula $\theta(a, S)$ in the language $\mathscr{L}_S$ we are able to translate it to an equivalent formulation $S(F_{\theta(x)}, a)$

where $\llcorner F_{\theta(x)} \lrcorner$ is a formula in the language $\mathscr{L}_A$. In other words, we are able to 'move' the instances of satisfaction 'outside' the formula. From this, if $\exists x \theta(x, S)$, then we can use propositional logic to show there is some assignment $b : x \to c$ such that $S(F_{\theta(x)}, b)$ and for all assignments $b' : x \to c'$ where $c' < c$ are such that $\neg S(F_{\theta(x)}, b')$. In other words, there is a least c such that $\theta(c, S)$ and the least number principle holds Δ_0 formulas of $\mathscr{L}_S$, which is equivalent to induction in $\mathscr{L}_S$ [4, Theorem 4]. $\square$

We shall explore the proof method and make these ideas more rigorous in Section 4.2. For the complete proof of this theorem the reader is referred to *Deflationary Truth and Pathologies* [4, Pages 9–11].
We see that whilst there are numerous different axiom schemas which remove pathological cases from satisfaction classes, these theories are very strong. We thus consider some weaker approaches with original considerations.

E.5.3 Adding Additional Tarski Conditions

An alternative method to restricting the pathologous examples in the theory of satisfaction classes is to add additional Tarski conditions that a satisfaction class must satisfy. In this section we consider an alternative way of considering $\neg$ in first order logic and how this affects the theory of satisfaction classes.

E.5.3.1 The $f_\neg$ Tarski Axiom

Rather than defining $\neg$ in the standard way, we can instead consider it as a function which, for clarity, we shall denote $f_\neg$. By introducing a function in this way, we are able to add a new Tarski condition that satisfaction classes must satisfy that appears not to be a consequence of the previous axioms. We first set out how this new definition of negation, $f_\neg$, behaves.

Definition E.5.11 ($f_\neg$)**.** *We define the function $f_\neg$ over all sentences of our language $\mathscr{L}$. We suppose for simplicity, without loss of generality, that $\mathscr{L}$ is purely relational. For each relation R in $\mathscr{L}$ we add a new relation symbol $\not R$. We also add a new symbol $\neq$. Let σ be an $\mathscr{L}$-sentence, then:*

- *If σ is of the form $t = s$, then $f_\neg(\sigma)$ is $t \neq s$.*
- *If σ is of the form $t \neq s$, then $f_\neg(\sigma)$ is $t = s$.*
- *If σ is of the form $R(u, v, \ldots)$, then $f_\neg(\sigma)$ is $\not R(u, v, \ldots)$.*
- *If σ is of the form $\not R(u, v, \ldots)$, then $f_\neg(\sigma)$ is $R(u, v, \ldots)$.*
- *If σ is of the form $\alpha \wedge \beta$, then $f_\neg(\sigma)$ is $f_\neg(\alpha) \vee f_\neg(\beta)$.*
- *If σ is of the form $\alpha \vee \beta$, then $f_\neg(\sigma)$ is $f_\neg(\alpha) \wedge f_\neg(\beta)$.*
- *If σ is of the form $\exists x\theta(x)$, then $f_\neg(\sigma) = \forall x f_\neg(\theta(x))$.*
- *If σ is of the form $\forall x\theta(x)$, then $f_\neg(\sigma) = \exists x f_\neg(\theta(x))$ [1, Page 871].*

It is an easy exercise to verify that $f_\neg$ behaves in the way we expect $\neg$ to behave. We can now consider a new axiom for satisfaction classes, which says that $S(\sigma)$ if and only if $\neg S(f_\neg(\sigma))$.

Definition E.5.12 ($f_\neg$ Tarski Axiom)**.** *We define the $f_\neg$ Tarksi Condition to be the sentence:*

8) For all $\mathscr{L}_A$-formulas θ : $S(\ulcorner\theta\urcorner, a) \leftrightarrow \neg S(\ulcorner f_\neg(\theta)\urcorner, a)$

This is an intuitively attractive axiom to have and is not a consequence of the other Tarski axioms.

Theorem E.5.13. *Consider a model $M \models \mathrm{PA}$ which is countable and recursively saturated. Then M has a full satisfaction class S satisfying Tarski Conditions 1)-7) and there is a sentence of $\mathscr{L}_A$, σ, such that $(M, S) \models S(\sigma)$ and $(M, S) \models S(f_\neg(\sigma))$.*

Proof Sketch. By Theorem E.5.1 that the sentence $(0 = 1 \vee (0 = 1 \vee (\cdots \vee (0 = 1) \cdots)))$ with a nonstandard number of disjuncts is consistent in M-Logic. By the soundness theorem it suffices to show that the set $\{\bigvee_t 0 = 1, \bigwedge_t 0 \neq 1\}$ is consistent in M-Logic. This is the case, since after finitely many provability rules of M-Logic we shall not reach $\vdash_M \sigma, \neg\sigma$ since not all of the disjuncts in the first sentence will be 'seen' by M-Logic, so M-Logic cannot tell that the two statements are contradictory. □

This additional Tarski condition is an easy and seemingly acceptable way of strengthening the theory of satisfaction classes. It is an open question as to the relative strength of this theory.

Question E.5.14. *What is the provability strength of the theory of satisfaction classes with the additional $f_\neg$ Tarski condition?*

There is one further method of attempting to remove pathological examples from the theory of satisfaction classes, which is seemingly the weakest idea. This approach ensures that certain sentences of a satisfaction class are always satisfied.

E.5.4 Specifying Certain Sentences as True

Some satisfaction classes will tell us that all sentences that we would like to be true, of a certain type, are true. When constructing a satisfaction class through the completeness of M-Logic, we are able to adapt the construction to ensure that certain sentences are always added to the construction. This method is described for collection in Section 2, but for now we shall explore some general results on what effect this has. To start, we provide some general notation .

Notation E.5.15. *Let M be a model such that $M \models \mathrm{PA}$ and S be a satisfaction class over M. For a set of $\mathscr{L}_A$ formulas $\Gamma = \{\gamma_1, \gamma_2, \ldots\}$ we write $S \models \Gamma$ to mean that $(M, S) \models S(\gamma_i)$ for each $\gamma_i \in \Gamma$.*

A natural set of sentences to ensure that a satisfaction class makes true are all the induction axioms. We shall denote this by $S \models \mathrm{IND}$. It is interesting to look at whether this behaves in a similar manner to induction without satisfaction. It is a standard theorem of the induction axioms that they are equivalent to the least number principle (LNP). We prove below, following the proof of induction impliying LNP in PA [12, Page 51], that this holds for satisfaction classes as well and making the least number principle true is equivalent to making induction true.

Theorem E.5.16. *Let $(M, S) \models \mathrm{PA}$ with a full satisfaction S. We have that $S \models \mathrm{IND}$ if and only if $S \models \mathrm{LNP}$.*

Proof. Firstly, suppose that $S \models \mathrm{IND}$ and $\theta(x)$ is an $\mathscr{L}_A$-formula. We require that: $(M, S) \models S(\ulcorner \exists x \theta(x) \to \exists x[\theta(x) \wedge \forall y{<}x \neg\theta(y)]\urcorner, a)$ which rephrases to $(M,S) \vDash \neg S(\ulcorner \exists x\theta(x), a) \vee S(\exists x[\theta(x) \wedge \forall y{<}x\neg\theta(y)]\urcorner, a)$. Suppose that $(M,S) \vDash \neg S(\ulcorner \exists x[\theta(x) \wedge \forall y{<}x\neg\theta(y)]\urcorner, a)$. Then, we have that:

$$(M,S) \vDash S(\ulcorner \forall x[\neg\theta(x) \vee \exists y{<}x\theta(y)]\urcorner, a).$$

Therefore, $(M,S) \vDash S(\ulcorner \forall x[\theta(x) \rightarrow \exists y{<}x\theta(y)]\urcorner, a)$. Thus, by contrapositive, we have that $(M,S) \vDash S(\ulcorner \forall x[\forall y{<}x\neg\theta(y) \rightarrow \neg\theta(x)]\urcorner, a)$. We define the formula $\phi(x)$ by $\forall z[z{<}x \rightarrow \neg\theta(z)]$. Clearly $(M,S) \vDash S(\ulcorner \phi(0)\urcorner, a)$ and from the above we see that $(M,S) \vDash S(\ulcorner \forall x[\phi(x) \rightarrow \phi(x+1)]\urcorner, a)$. Therefore, since $S \vDash \text{IND}$ we have that $(M,S) \vDash S(\ulcorner \forall x\phi(x)\urcorner, a)$, which entails that $(M,S) \vDash S(\ulcorner \forall x\neg\theta(x)\urcorner, a)$ which by Tarski condition 5) tells us that $(M,S) \vDash \neg S(\ulcorner \exists x\theta(x)\urcorner, a)$.

Now suppose that $S \vDash \text{LNP}$ and that $(M,S) \vDash S(\ulcorner \theta(0) \wedge \forall x[\theta(x) \rightarrow \theta(x+1)]\urcorner, a)$ for any ${}^\star\mathscr{L}_A$-formula $\theta(x)$. Suppose for contradiction that $(M,S) \vDash S(\ulcorner \exists x[\neg\theta(x)]\urcorner, a)$. Then, since $S \vDash \text{LNP}$ we have that $(M,S) \vDash S(\ulcorner \forall y{<}x\theta(y)\urcorner, a[{}^x/_c])$ for some constant $c \in M$ such that:

$$(M,S) \vDash S(\ulcorner \neg\theta(x)\urcorner, a[{}^x/_c]).$$

This tells us that $c > 0$ and that $(M,S) \vDash S(\ulcorner \theta(x)\urcorner, a[{}^x/_{c-1}])$ and so therefore by our original hypothesis $(M,S) \vDash S(\ulcorner \theta(x)\urcorner, a[{}^x/_c])$ which contradicts Tarski condition 5). □

In building a satisfaction class S which satisfies all induction axioms we are able to remove any pathological induction axioms. However, unlike the thoery of satisfaction classes including induction in $\mathscr{L}_S$ this is not as strong and does not give us any induction on formulas involving the satisfaction predicate and other formulas and connectives as well. It is an open question as to how much induction this approach provides and how comparatively strong our theory is.

Question E.5.17. *How strong is the theory of S is a full satisfaction class where $S \models \mathrm{IND}$?*

We see here we have three different general strategies to improving the theory of satisfaction classes to remove pathological examples. We can add further axioms to $\mathscr{L}_S$, add further Tarski conditions and specify certain $\mathscr{L}_A$ sentences as always satisfied.

Appendix F

Open Questions and Conjectures

Conjecture 2.2.1. *There are models* $M \models \mathrm{PA}$ *with a full satisfaction class* S *such that* $S \models \mathrm{I}\Delta_0 + \mathrm{Coll}$ *and* $S \not\models \mathrm{IND}$.

Question 3.1.4. *Is there a model* $M \models \mathrm{PA}$ *with a satisfaction class* S *such that* $(M, S) \not\models \mathrm{Coll}(S)$ *and* $(M, S) \not\models \Delta_0 - \mathrm{PA}(S)$.

Question 3.1.5. *Is there a model* $M \models \mathrm{PA}$ *with a satisfaction class* S *such that* $(M, S) \models \mathrm{Coll}(S)$ *and* $(M, S) \not\models \Delta_0 - \mathrm{PA}(S)$.

Conjecture 3.2.4. *Let* $(M, S) \models \mathrm{PA} + S$ *is a full satisfaction class over* M*, where* $S \models \mathrm{IND}$*. Let* N *denote* $cl^{\mathbb{N}}_{Skolem}(\varnothing)$ *and define* $S_N \subseteq N \times N$ *by* $(N, S_N) \models S_N(\ulcorner\theta\urcorner, a)$ *if and only if* $(M, S) \models S(\ulcorner\theta\urcorner, a)$*. Then,* $(N, S_N) \not\models \mathrm{Coll}_{\Sigma_1}(S)$.

Conjecture 4.1.3. *We can adapt Visser-Enayat's Theorem to obtain an end extension, rather than just an extension. In other words, given* $(M, F_M, S_M) \models$

$\mathrm{PA} + S_M$ *is an* F_M *satisfaction class there exist* N, F_N *and* S_N *such that* $(M, F_M, S_M) \subseteq_e (N, F_N, S_N)$.

Question 4.1.6. *Is there an analogue to the MRDP Theorem in the language* $\mathscr{L}_S$*? In other words, is it the case that* $\exists_1(S) = \Sigma_1(S)$*?*

Question 4.2.4. *What relative consistency strength does the theory* $\mathrm{PA} + S$ *is a full satisfaction class* $+S$ *satisfies Tarski Conditions 9) and 10) have over the theory* $\mathrm{PA} + S$ *is a full satisfaction class?*

Conjecture 4.3.2. *We can construct satisfaction classes using the Visser-Enayat method which satisfy Tarski Conditions* $9)^\star_n$ *and* $10)^\star_n$ *for each* $n \in \mathbb{N}$.

Conjecture 4.3.5. *The theory of* $\mathrm{PA} + S$ *is a full codable satisfaction class has the same provability strength as* $\Delta_0 - \mathrm{PA}(S)$.

Question 4.3.11. *Does a structure* (M, F, S) *where* $M \models \mathrm{PA}$ *and* S *is an* F*-satisfaction class and* F *is a set of stratified formulas satisfy Lachlan's Thorem?*

Question 4.3.12. *Given* M, F *where* $M \models \mathrm{PA}$ *and* F *is a set of stratified formulas, does there exist* S *such that* S *is an* F*-satisfaction class?*

Question 5.5. *Can the nonstandard language of* $\mathscr{L}_A$ *be restricted in such a way so that it loses no expressive content and so that there exists a deflationary truth predicate for* PA *over all sentences of this restricted language?*

Question D.2.17. *Are there any further connections between* $\mathrm{Coll}\bigwedge$, $\mathrm{Coll}\bigvee$, *tallness and short recursive saturation?*

Question E.5.14. *What is the provability strength of the theory of satisfaction classes with the additional* $f_\neg$ *Tarski condition?*

Question E.5.17. *How strong is the theory of* S *is a full satisfaction class where* $S \models \mathrm{IND}$*?*

Bibliography

[1] *Handbook of mathematical logic.* North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. Edited by Jon Barwise, With the cooperation of H. J. Keisler, K. Kunen, Y. N. Moschovakis and A. S. Troelstra, Studies in Logic and the Foundations of Mathematics, Vol. 90.

[2] A. J. Ayer. The criterion of truth. *Analysis*, 3(1/2):28–31, 1935.

[3] Chantal Berline, Kenneth McAloon, and Jean-Pierre Ressayre, editors. *Model theory and arithmetic*, volume 890 of *Lecture Notes in Mathematics*. Springer, Berlin, 1981.

[4] Cezary Cieśliński. Deflationary truth and pathologies. *J. Philos. Logic*, 39(3):325–337, 2010.

[5] Ali Enayat and Albert Visser. New constructions of satisfaction classes. 2013.

[6] Gottlob Frege. The thought: A logical inquiry. *Mind*, 65(259):289–311, 1956.

[7] Wilfrid Hodges. *Building models by games*, volume 2 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1985.

[8] Paul Horwich. *Truth*. Clarendon Press, 1998.

[9] Carol R. Karp. *Languages with expressions of infinite length.* North–Holland Publishing Co., Amsterdam, 1964.

[10] Richard Kaye. *Models of Peano arithmetic*, volume 15 of *Oxford Logic Guides*. The Clarendon Press, Oxford University Press, New York, 1991. Oxford Science Publications.

[11] Richard Kaye. *The mathematics of logic*. Cambridge University Press, Cambridge, 2007. A guide to completeness theorems and their applications.

[12] Richard Kaye. Models of Peano arithmetic. Second Edition, please contact the author for a copy, 2012.

[13] Jeffrey Ketland. Deflationism and tarski's paradise. *Mind*, 108(429):69–94, 1999.

[14] 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.

[15] A. H. Lachlan. Full satisfaction classes and recursive saturation. *Canad. Math. Bull.*, 24(3):295–297, 1981.

[16] W. V. Quine. *Philosophy of Logic*. Harvard University Press, 1986.

[17] Abraham Robinson. On languages which are based on non-standard arithmetic. *Nagoya Math. J.*, 22:83–117, 1963.

[18] Richard Schantz. Was tarski a deflationist? *Logic and Logical Philosophy*, 6:157, 2014.

[19] Stuart Thomas Smith. *NONSTANDARD SYNTAX AND SEMANTICS AND FULL SATISFACTION CLASSES FOR MODELS OF ARITHMETIC*. ProQuest LLC, Ann Arbor, MI, 1984. Thesis (Ph.D.)–Yale University.

[20] Alfred Tarski. *Logic, semantics, metamathematics. Papers from 1923 to 1938*. Oxford at the Clarendon Press, 1956. Translated by J. H. Woodger.