Cyclic Proofs for Theories of Truth

Thesis for the Master's degree in Logic, 30 credits

Author:
Bente Gortworst

Supervisors:
Graham Leigh
Gianluca Curzi

UNIVERSITY OF GOTHENBURG
Department of Philosophy, Linguistics and Theory of Science
Defended on 31st May 2024

Of truth we know nothing,
for truth is in a well.

Democritus

Abstract

Formal theories of truth aim to help us understand the notion of truth and its role in logic and mathematics. The field has come forth from Tarski's effort to give a formal truth definition, and has since seen significant developments. One of the guiding theories on mathematical truth is Kripke's semantic theory which partially models the truth predicate as a minimal fixed point. Kripke's theory has motivated several axiomatic theories that capture this truth notion. Among them the most popular seems to be the theory Kripke-Feferman (KF). Investigations into the proof theory of KF have been ongoing and inform our understanding of the relation between truth and provability in mathematical logic.

In this thesis we contribute to these investigations through the lens of cyclic proofs, a type of formal proof that allows for finite proof-theoretic representations of possibly infinite derivations through cyclic reasoning. This type of proof has received much attention in recent years due to its practical applicability, and has been analyzed in the context of several logical theories, but has not yet been applied to theories of truth. Therefore we bring together the streams of research by developing a cyclic proof system for the theory KF. We also demonstrate we are able to further expand this system into cyclic systems for μKF (a version of KF where truth and falsity are taken to be minimal fixed points), and νKF (which takes truth and falsity to be maximal fixed points). We show that all three systems are sound with respect to their semantic construction and that we can simulate inductive reasoning through cyclic rules in these systems.

Acknowledgements

Firstly, I would like to sincerely thank my supervisors Graham and Gianluca for all their help and guidance throughout this process. This thesis would not be even close to presentable without their comments and our many meetings. Thank you for making me a better researcher and logician.

I would also like to thank everyone I met at the Logic group: all the teachers, PhD students and other master's students, past and present. I have greatly enjoyed my time with you these two years, and I am grateful to have been part of such a wonderful community. Specifically, I want to thank Tosia, with whom I have been on this academic journey for many years, and who really is the reason I came to Gothenburg. And I want to thank Kirsten and Erik, my writing companions, my study buddies, for all the coffee breaks and crisis meetings.

Lastly I want to thank all my other friends and family who will never read this thesis but who I can always count on for support. And thank you to Ata, for making me feel at peace.

Contents

Chapter 1

Introduction

1.1 Truth, proof and paradox

Defining truth is an endeavour that has occupied philosophers and mathematicians for centuries, with the discussion going back to ancient Greece. We can find traces of it in texts as early as Aristotle's *Metaphysics*[1]. Ever since, philosophers have attempted to give definitions of truth in terms of, for instance, coherence, correspondence and pragmatism [HL22]. Often these theories would take truth to be a substantial object, but they do not give a *formally* precise definition of truth in itself.

Paradoxes have played a big role in the discussion around truth, specifically the question of how to cope with their existence in our language. Take for instance the so-called *liar sentence*, denoted by L, that asserts its own falsity (it says of itself: "I am not true"); then L is true iff L is false, which is inherently contradictory. Though antinomies like this have been a concern since the antiquity (the first formulation of the liar paradox has been ascribed to the ancient Greek philosopher Eubulides [Tar69, p. 66]), how they should be resolved in formal definitions of truth is still contentious. On the side of mathematics, truth has often been equated with provability: it was assumed that by constructing a mathematical proof we demonstrate the truth of a statement [Tar31, p. 169], and that any true mathematical statement can be established by a rigorous argument. However, the notions of both mathematical proof and truth remained rather imprecise for a long time, making it difficult to formally connect the two.

In the 1930s, Alfred Tarski made an effort to define the concept of mathematical truth in formal logic [Tar31]. This gave rise to a new programme whose goal was to further develop formal theories of truth. These modern truth theories aim to either describe mathematical truth through semantic definitions or through axiomatic systems, where truth boils down to a list or axioms that must be satisfied in order for a statement to be considered true

[1] Here Aristotle said of truth, as quoted by Tarski: "To say of what it is that it is not, or of what is not that it is, is false, while to say of what it is that it is, or of what is not that it is not, is true" [Tar69, p. 63]

[HL22]. Formal theories of truth, either semantic or axiomatic, can shed light on how we view and use truth implicitly in logic, philosophy and mathematics.

Around the same time as the development of Tarski's truth definition, mathematician and logician David Hilbert initiated a programme aimed at studying the notion of a mathematical proof through tools of formal logic. This was the first substantial set up for the branch of mathematical logic we now know as proof theory [Mur20, p. 12]. Both the proof-theoretic programme initiated by Hilbert and Tarki's foundational research on formal definitions of truth revealed that the notion of truth and that of provability do not always coincide in mathematics. On the one hand, this is demonstrated by Gödel's incompleteness theorems, which showed that there are true arithmetical sentences that cannot be proved within the formal theory of arithmetic, [Ken20]. On the other hand, Tarski's undefinability theorem demonstrated that the set of statements that are true does not coincide with the set of formally provable statements, since the former cannot be defined in arithmetic, while the latter can [Tar69, p. 76]. These developments highlight a fundamental gap between the notion of truth and that of provability, which has since motivated a systematic proof-theoretic investigation of formal theories of truth.

The current thesis builds on ongoing research in both formal theories of truth and proof theory. For the latter, we will specifically be looking into a subdivision of proof theory which has been gaining attention over recent years, that concerns cyclic proofs. These are proof systems that allow for possibly infinite derivations by introducing a cyclic argumentation style, which replaces the usual inductive reasoning more commonly used in formal proofs. Some of the first literature informing modern investigations into cyclic proof theory is attributed to Brotherston [Bro05] and Simpson [Sim17]. In recent years especially, the field has seen significant growth. Cyclic proofs have proven themselves as a useful alternative to the more traditional proofs relying on induction, especially in applications for automated reasoning (cf. [BGP12], [TB20], [TU22], [BDP11]). Outside of practical applications, cyclic proof theory has been investigated extensively in the context of various logical theories, but not yet in the context of theories of truth. Therefore this thesis adds a novel perspective to recent and promising research into cyclic proofs, by investigating truth theories through the lens of cyclic proofs. Our ultimate aim is to establish a cyclic proof system for the axiomatic truth theory known as Kripke-Feferman. We also construct auxiliary systems for some theories related to KF, namely Burgess's theory μKF [Bur14] and the theory νKF, which we develop in this thesis. In doing so, we contribute to our understanding of how cyclic proofs for specific theories of truth should look and what properties of truth contribute to the presentation.

1.2 Motivations

The research in this thesis is interesting for two main reasons. On the side of formal theories of truth, it analyzes the interactions between syntactic representations of axiomatizations of truth and the semantic constructions of truth in the setting of cyclic proof theory. As far as the author is aware, there is no prior literature attempting to bridge the gap between theories of truth and cyclic proof theory, and therefore this thesis is the first work in this direction. Questioning the connections between these streams of research allows us to gain a better understanding of how we axiomatize and reason about truth in logic. On the side of cyclic proof theory, we shall see how the properties of truth theories,concerning how fixed points are repesented in the theory, affect the global proof-theoretic structure of cyclic proofs, as well as what conditions ought to be placed on cyclic systems, advancing our understanding of cyclic proofs in the setting of such theories.

The choice of the theory KF as the main focus of this thesis is motivated by the semantic theory of Kripke at its foundation, which defines truth in terms of a fixed point construction. Since cyclic proofs have established themselves as an amenable framework for modelling fixed points, we determined that KF is particularly well-suited for a cyclic proof-theoretic analysis. Alternative theories of truths were initially considered as well for this research, but ultimately disregarded. However, analyzing these other theories in the setting of cyclic proofs could be interesting as a potential future research direction.

1.3 Structure

We now briefly outline the structure of this thesis. We will first establish the necessary background information on semantic and axiomatic theories of truth in Chapter 2, where the focus will be on Kripke's truth definition and axiomatizations of it. Then we will provide background information on cyclic proof theory and the proof system of cyclic Peano arithmetic in Chapter 3. We will build on this for the main findings of the thesis, which start with our development of a proof system consisting of a finitary sequent calculus for the theory Kripke-Feferman (KF) and related theories μKF and νKF in Chapter 4. These will be further developed into cyclic proof systems for the mentioned truth theories in Chapter 5. We will show that the resulting cyclic proof systems are sound, and that they can embed the inductive reasoning of their non-cyclic counterparts through cyclic rules. The implications of these contributions will be discussed in the Discussion, Chapter 6, where we also note some possible further directions to extend the current research.

Notational Conventions We let $\mathcal{L}$ be the language of Peano arithmetic (PA), including the induction scheme. We say a theory is formulated in the language $\mathcal{L}_T$ when it is formulated in $\mathcal{L}$ with PA's induction scheme extended to possibly include induction over the *truth predicate* T [Hal14, p. 89]. We call this extension of Peano arithmetic PA_T. The language $\mathcal{L}_{T,F}$ denotes the further expansion of this language to include induction over the *falsity predicate* F. The language $\mathcal{L}^+_{T,F}$ restricts the occurrence of T and F to only *positive* instances; meaning these predicates cannot occur in the scope of any negations. The negation symbol in this language is only allowed to occur in front of some atomic formula of $\mathcal{L}$ [Lei16, p. 575].

In this thesis, when formulating theories of truth we will usually be speaking either in $\mathcal{L}_T$ or $\mathcal{L}^+_{T,F}$. Expressions in these languages can be coded in the natural numbers through Gödel coding. We will briefly set out some of the ways this coding is described in the remainder of this thesis.

The Gödel codes of terms in the language of a truth theory, commonly denoted as $\ulcorner t \urcorner$, is generally abbreviated by giving the term with a dot underneath, so we write $\d{t}$. Likewise the code of a formula A is denoted as $\d{A}$. Furthermore, predicates that have a dot underneath them ($\d{=}, \d{\neg}, \d{\wedge}, \d{\forall}$, etc.) represent the terms obtained from operations on Gödel codes. So, for instance, $s \d{\wedge} t := \d{s} \wedge \d{t}$. The statement $T\d{T}x$ abbreviates the statement $\forall y.(y = (\d{T}[x/\d{x}]) \rightarrow \d{T}y)$, likewise when replacing T with F. Operations of the form $T\d{t}$, $F\d{t}$, $T\d{A}$ and $F\d{A}$ may, with some abuse of notation, be simplified to Tt, Ft, TA and FA. Thus in such cases we tend to not distinguish between expressions and their Gödel codes.

We introduce the notation $\mathrm{Sent}_{\mathcal{L}}(A)$ to denote that A is the code of a sentence of $\mathcal{L}$. A sentence is defined as usual, meaning a formula A is a sentence iff it contains no free variables. An expression of the form t° represents the value of the *numeral* associated with the closed term t. A closed term is a term containing no variables. We write $\mathrm{FV}(A)$ to denote the set of free variables in A.

Chapter 2

Formal Theories of Truth

Since Tarski gave a formal definition of truth in his paper *The concept of truth in formalized languages*, the study of formal theories of truth has become an important research area within mathematical logic. These modern truth theories can be roughly divided into two categories which take a different approach to truth: *semantic theories*, and *axiomatic theories*. Semantic theories aim to describe truth for a formal language by giving a truth predicate and then defining this predicate in some stronger "metalanguage". Axiomatic theories, on the other hand, instead describe the behaviour of this truth predicate with a set of axioms operating on the same level as the formal language. As will become apparent in this chapter, the two approaches are inherently intertwined, with axiomatic theories usually building on the foundations of a semantic theory.

2.1 Semantic Truth Theories

2.1.1 Tarski's semantic truth definition

Tarski's definition of truth in formal languages is constructed in terms of the satisfaction of sentences in a given language, the *object language*, and formulates truth in a so-called *metalanguage* [Tar69, p. 68]. This metalanguage contains the object language, but is also able to talk about the object language. To this end, the metalanguage is extended with some predicate T which expresses the truth value of a sentence in the object language. The property of a sentence being true can be formalized as what he calls *Convention T*, originally defined in the following way:

Definition 2.1 (Convention T, [Tar31])**.** A formal definition of the class of all true sentences T will be called adequate if it has the following consequences:

1. all sentences which are obtained from the expression "$x \in T$ if and only if p", where:
    - the symbol x represents any sentence of the object language, and

- the symbol p represents the translation of this sentence into the metalanguage,

2. the sentence "for any x, if $x \in T$ then $x \in S$", where S is the class of all sentences.

The corresponding semantic truth theory recursively decides the truth of sentences in a structure based on the satisfaction of Convention T. According to his definition, if we, for instance, take the following sentence:

"Snow is white"

We can assert this sentence to be true in the object language if and only if the statement that snow is white belongs to the metalanguage.

Along with his truth definition, Tarski gave his theorem on the *undefinability* of truth, which has been greatly influential for formal truth theories since. The undefinability theorem tells us that, in the framework of classical logic, a complete definition of truth in an object language cannot be given in that language itself, but rather has to be formulated in some stronger metalanguage [Tar31, p. 254]. This follows from Convention T, by showing that if there is some arithmetical formula that defines truth in the language of arithmetic, then there is some arithmetical sentence (namely the liar sentence L) that asserts of itself that it is not an arithmetical truth. By Convention T, we can then formulate a sentence $L \leftrightarrow T\underset{\cdot}{L}$, which leads to an inconsistency [Hor11, p. 35]. Therefore Tarski demonstrates that we cannot have any classical model $\mathfrak{M}$ such that for *all* sentences A of $\mathcal{L}_T$ we have $\mathfrak{M} \models A$ iff $\mathfrak{M} \models T\underset{\cdot}{A}$. In fact, we can only define truth within an object language for a fraction of the language that escapes the liar paradox. Thus the set of all true sentences in arithmetic is not definable in the language of arithmetic. However, since it is possible to define the set of provable sentences in the language of arithmetic [Tar31, p. 254], Tarski's undefinability theorem highlights that truth cannot be equated with provability in classical logic.

Tarski's formal definition shifts away from the previous trends describing truth in terms of substantial notions like correspondence and coherence. It instead takes truth as functional, emphasising questions about how it can be used in mathematics and logic, rather than questioning what its essence is (a rather vague and informal question) [Hor11, p. 16]. He thereby lays the foundation for the modern theories of truth that build upon these concepts. Tarski's approach is semantic. He decided not to pursue the axiomatic approach, because, as he noted, a truth axiomatization based on sentences that satisfy Convention T alone would not be strong enough to say anything useful about its object language [Hal14, p. 18]. Furthermore he argued that any additional axioms that could be added to make a stronger theory would have an arbitrary character.

Tarski's theory only defines truth for a small fragment of a language, namely the sentences which do not contain any truth predicates in themselves [Tar31, p. 165]. By limiting ourselves to this fragment, we circumvent the paradox created by the liar sentence, as demonstrated in the undefinability theorem. Several theories following Tarski's example adhere to this notion, and give a truth predicate that can only be applied to truth-free sentences. These are the *typed* theories. On the other side there are theories which do allow for a self-applicable truth predicate, *type-free* theories [HL22].

2.1.2 Kripke's semantic truth definition

One of the first and most popular proposals for a type-free approach to defining truth is Kripke's [Kri75]. He gives an outline of a semantic theory where the truth predicate T is self-applicable but only partially interpreted, therefore asserting that it may be possible for a sentence to be neither true nor false. In Kripke's semantic theory of truth, models take the form $(\mathfrak{M}, S_1, S_2)$, where S_1 and S_2 are subsets of the domain of $\mathfrak{M}$ called, respectively, its *extension* and *anti-extension.* The truth value of sentences in in the language of $\mathfrak{M}$ is then established in terms of Kleene's strong three valued logic [Kri75, p. 700]. When $\mathfrak{M}$ is taken to be the standard model of Peano arithmetic $\mathbb{N}$, the *strong Kleene valuation scheme* is defined as follows (see also [HH06, p. 680], and [McG90, p. 88]):[1]

Definition 2.2. For the structure $\mathbb{N}$, and $S_1, S_2 \subseteq \omega$, $(\mathbb{N}, S_1, S_2) \vDash_{sk} A$ is inductively defined as follows, with s, t closed terms and A, B $\mathcal{L}_T$-sentences:

1. $(\mathbb{N}, S_1, S_2) \vDash_{sk} s = t$ iff s and t's values coincide;
2. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg s = t$ iff s and t's values do not coincide;
3. $(\mathbb{N}, S_1, S_2) \vDash_{sk} Tt$ iff t is a closed term whose value is an $\mathcal{L}_T$-sentence in S_1;
4. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg Tt$ iff t is a closed term whose value is an $\mathcal{L}_T$-sentence in S_2, or whose value is not an $\mathcal{L}_T$-sentence;
5. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg\neg A$ iff $(\mathbb{N}, S_1, S_2) \vDash_{sk} A$;
6. $(\mathbb{N}, S_1, S_2) \vDash_{sk} A \wedge B$ iff $(\mathbb{N}, S_1, S_2) \vDash_{sk} A$ and $(\mathbb{N}, S_1, S_2) \vDash_{sk} B$;
7. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg(A \wedge B)$ iff $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A$ or $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg B$;
8. $(\mathbb{N}, S_1, S_2) \vDash_{sk} A \vee B$ iff $(\mathbb{N}, S_1, S_2) \vDash_{sk} A$ or $(\mathbb{N}, S_1, S_2) \vDash_{sk} B$;
9. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg(A \vee B)$ iff $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A$ and $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg B$;
10. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \forall x.A(x)$ iff for all $n \in \omega$ $(\mathbb{N}, S_1, S_2) \vDash_{sk} A(\bar{n})$;
11. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg\forall x.A(x)$ iff there exists some $n \in \omega$ such that $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A(\bar{n})$;

[1] We present here a version of the strong Kleene scheme where the extension and anti-extension are allowed to overlap, thus creating possibly both truth value gaps and gluts. While Kripke allowed for truth value gaps in his construction, he did not allow truth value gluts; there could never be sentences that are in both the extension and the anti-extension [Kri75]. This requirement is discarded here to be in line with the axiomatization of this theory presented later.

12. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \exists x.A(x)$ iff there exists some $n \in \omega$ such that $(\mathbb{N}, S_1, S_2) \vDash_{sk} A(\bar{n})$;

13. $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg\exists x.A(x)$ iff for all $n \in \omega$ $(\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A(\bar{n})$

Note that the resulting models are non-classical, because the strong Kleene scheme does not satisfy the compositional axiom $\forall t.(T \underset{\cdot}{\neg} t \leftrightarrow \neg Tt)$ [HL22]. Hence we cannot claim in this system that saying a sentence being not true is equivalent to saying its negation is true. This is what allows us to avoid the liar paradox.

In Kripke's construction, we define a model for the language $\mathcal{L}_T$ by iteratively building the extension and the anti-extension as a partial expansion of the standard model of arithmetic $\mathbb{N}$. To this end, we construct the following function Φ on pairs of sets (S_1, S_2) of natural numbers, called the *Kripke jump operator*. We can define it as:

$$\Phi(S_1, S_2) = (\{A \in \mathcal{L}_T : (\mathbb{N}, S_1, S_2) \vDash_{sk} A\}, \quad \overline{\mathrm{Sent}}_{\mathcal{L}_T} \cup \{A \in \mathcal{L}_T : (\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A\})$$

where $\overline{\mathrm{Sent}}$ denotes the set of all numbers that are not codes of sentences in $\mathcal{L}_T$. The operator Φ returns the set of sentences valid in S_1 and S_2, as well as the set of non-sentences and false sentences in S_1 and S_2 ([HH06, p. 681]).

Fact 2.3. ([Hor11, p. 119]) Φ is a monotone operator, in the sense that if $S_1 \subseteq S_1'$, $S_2 \subseteq S_2'$, then
$\{A \in \mathcal{L}_T : (\mathbb{N}, S_1, S_2) \vDash_{sk} A\} \subseteq \{A \in \mathcal{L}_T : (\mathbb{N}, S_1', S_2') \vDash_{sk} A\}$, and
$\overline{\mathrm{Sent}} \cup \{A \in \mathcal{L}_T : (\mathbb{N}, S_1, S_2) \vDash_{sk} \neg A\}) \subseteq \overline{\mathrm{Sent}} \cup \{A \in \mathcal{L}_T : (\mathbb{N}, S_1', S_2') \vDash_{sk} \neg A\}$.

By the Knaster-Tarski theorem[2] the monotonicity of Φ guarantees the existence of a least fixed point $(\mathcal{E}, \mathcal{A})$ for Φ. That is, there is an $\mathcal{E}$ and $\mathcal{A}$ such that $\Phi(\mathcal{E}, \mathcal{A}) = (\mathcal{E}, \mathcal{A})$ [HH06, p. 681]. Kripke's *least fixed point model*, denoted as $\mathfrak{K}$, is then defined as $\mathfrak{K} := (\mathbb{N}, \mathcal{E}, \mathcal{A})$.

We can also define the least fixed point model by way of a transfinite construction. We will denote models constructed in each iteration of this process as $\mathfrak{K}_\alpha$, where α denotes we are in the α-th stage of the transfinite construction. At stage 0, we have the model $\mathfrak{K}_0$ where both the extension and anti-extension are empty, so $\mathfrak{K}_0 := (\mathbb{N}, \varnothing, \varnothing)$. At the $\alpha + 1$-th iteration of the construction, we let $\mathfrak{K}_{\alpha+1} := (\mathbb{N}, \mathcal{E}_{\alpha+1}, \mathcal{A}_{\alpha+1})$ such that $\mathcal{E}_{\alpha+1}$ and $\mathcal{A}_{\alpha+1}$ are determined as the sentences that are made true, or respectively false, in $\mathfrak{K}_\alpha$. From the strong Kleene scheme it follows that any partial model $\mathfrak{K}_\alpha$ has the monotonicity property [Hor11, p. 119]. Formally:

[2] This theorem states that, for a complete lattice $\langle A, \leqslant\rangle$ and a monotone function $f : A \to A$ that is order-preserving with regard to $\leqslant$, then the set P of all fixed points of f is non-empty and $\langle P, \leqslant\rangle$ forms a complete lattice [Tar55, p. 286]. In particular this implies the existence of a least and greatest fixed point of f.

Definition 2.4 (Iterated construction, [Hor11]). For any ordinal α we define $\mathfrak{K}_\alpha := (\mathbb{N}, \mathcal{E}_\alpha, \mathcal{A}_\alpha)$ as follows:

$$\begin{array}{rclrcl} \mathcal{E}_0 & := & \varnothing & \mathcal{A}_0 & := & \varnothing \\ \mathcal{E}_{\alpha+1} & := & \{A \in \mathcal{L}_T : \mathfrak{K}_\alpha \models_{sk} A\} & \mathcal{A}_{\alpha+1} & := & \{A \in \mathcal{L}_T : \mathfrak{K}_\alpha \models_{sk} \neg A\} \\ \mathcal{E}_\lambda & := & \bigcup_{\alpha<\lambda} \mathcal{E}_\alpha & \mathcal{A}_\lambda & := & \bigcup_{\alpha<\lambda} \mathcal{A}_\alpha \end{array}$$

where λ is a limit ordinal.

At some ordinal stage θ, the process of adding new sentences to previously constructed models halts, as further iteration does not provide new information on the truth value of the sentences in the model. For this ordinal θ we have $\mathcal{E}_{\theta+1} = \mathcal{E}_\theta$ and $\mathcal{A}_{\theta+1} = \mathcal{A}_\theta$. We know that the least such ordinal, called the *closure ordinal*, is ω_1^{CK}: the first non-recursive ordinal (the first ordinal that cannot be represented through either natural numbers or a recursive function) [Hor11, p. 119-120]. In other words, at this stage we have obtained the least fixed point model, so $\mathfrak{K} := \mathfrak{K}_{\omega_1^{CK}}$. Thus, $\mathfrak{K}_\theta$ is a *fixed point model* for any $\theta \geqslant \omega_1^{CK}$. We have:

Theorem 2.5 ([Hor11]). *$\mathfrak{K}_\theta = \mathfrak{K}_{\theta+1}$ for each $\theta \geqslant \omega_1^{CK}$.*

Kripke's fixed point models have several good properties, not least of which is that the unrestricted *Tarski biconditionals*, that is, sentences of the form $A \leftrightarrow T\underset{\cdot}{A}$, hold in it:

Corollary 2.6 ([Hor11]). *For all sentences A in $\mathcal{L}_T$:*

$$\mathfrak{K}_\theta \models_{sk} A \text{ if and only if } \mathfrak{K}_\theta \models_{sk} T\underset{\cdot}{A}$$

where $\mathfrak{K}_\theta$ is any fixed point model of the Kripke construction.

Proof. The left-to-right direction follows from Theorem 2.5, noticing that $\mathfrak{K}_\theta \models_{sk} A$ implies $\mathfrak{K}_{\theta+1} \models_{sk} T\underset{\cdot}{A}$ by construction. For the other direction, suppose $\mathfrak{K}_\theta \models_{sk} T\underset{\cdot}{A}$. Then, for some ordinal $\beta < \theta$, $\mathfrak{K}_\beta \models_{sk} A$. By monotonicity, $\mathfrak{K}_\theta \models_{sk} A$. □

It follows from Corollary 2.6, that Tarski's undefinability theorem does not hold in the Kripke fixed point models, since we can assert the truth of any statement A in our object language within the same language [Hor11, p. 121]. Note that, by the construction of this model there may be sentences that are neither in the extension nor the anti-extension of the model. Thus T is interpreted as a partial predicate in the Kripke construction, meaing some sentences can have an undetermined truth value. The liar sentence is one such sentence. Hence the liar paradox leading to the undefinability theorem is circumvented and this theorem does not hold in this context.

Alternative construction of the Kripke model

The above construction of Kripke's fixed point model can equivalently be given in terms of a positive inductive definition of truth. This is the construction given by for instance McGee in [McG90, p. 107-126] and Halbach in [Hal14, p. 188-191], where a *Kripke truth set* is defined inductively, making the strong Kleene scheme implicit. We give a version of the construction of Kripke truth set that also introduces a predicate F for falsity[3], as well as a Kripke falsity set. This restricts us to formulate the Kripke truth set over the language $\mathcal{L}^+_{T,F}$, with truth and falsity predicates occurring only positively.

Definition 2.7 (Kripke truth-falsity pair)**.** [4] A pair of sets $S, \bar{S}$ is a *Kripke truth-falsity pair* iff the set $S \subseteq \omega$ is a *Kripke truth set* and $\bar{S} \subseteq \omega$ is a *Kripke falsity set*.

S is defined as a *Kripke truth set* iff it satisfies that $n \in S$ iff n is a $\mathcal{L}^+_{T,F}$-sentence and one of the following is true:

1. n is $(s = t)$ and s and t are closed terms whose values coincide;
2. n is $\neg(s = t)$ and s and t are closed terms whose values do not coincide;
3. n is $(A \wedge B)$ and $A \in S$ and $B \in S$;
4. n is $(A \vee B)$ and $A \in S$ or $B \in S$;
5. n is $(\forall x.A(x))$ and for all closed terms t, $A[t/x] \in S$;
6. n is $(\exists x.A(x))$ and there is some closed term t such that $A[t/x] \in S$;
7. n is Tt and t is a closed term whose value is in S;
8. n is Tt and t is a closed term whose value is in $\bar{S}$;

$\bar{S}$ is defined as a *Kripke falsity set* iff it satisfies that $n \in S$ iff n is a $\mathcal{L}^+_{T,F}$-sentence and one of the following is true:

1. n is $(s = t)$ and s and t are closed terms whose values do not coincide;
2. n is $\neg(s = t)$ and s and t are closed terms whose values coincide;
3. n is $(A \wedge B)$ and $A \in \bar{S}$ or $B \in \bar{S}$;
4. n is $(A \vee B)$ and $A \in \bar{S}$ and $B \in \bar{S}$;
5. n is $(\forall x.A(x))$ and there is some closed term t such that $A[t/x] \in \bar{S}$;

[3] Technically we can define F in terms of T, as $Ft := T\neg t$

[4] This definition is taken from [Hal14, p. 189], but adapted to also take into consideration a falsity predicate.

6. n is $(\exists x.A(x))$ and for all closed terms t, $A[t/x] \in \bar{S}$;

7. n is Tt and t is a closed term whose value is in $\bar{S}$, or the value of t is not an $\mathcal{L}^{+}_{T,F}$-sentence;

8. n is Tt and t is a closed term whose value is in S, or the value of t is not an $\mathcal{L}^{+}_{T,F}$-sentence;

From the Kripke truth-falsity pair S we can describe the operators

$$\Lambda(S) = \{n : \zeta(n, S)\}$$

$$\bar{\Lambda}(\bar{S}) = \overline{\mathrm{Sent}}_{\mathcal{L}^{+}_{T,F}} \cup \{n : \xi(n, \bar{S})\}$$

where $\zeta(n, S)$ is a formula expressing one of the clauses of the Kripke truth set S in Definition 2.7, where S occurs positively [Hal14, p. 190]. Likewise, $\xi(n, \bar{S})$ expresses one of the clauses of $\bar{S}$ in Definition 2.7, where $\bar{S}$ occurs positively. Thus we can define a model $\mathfrak{K}^{\alpha} := (\mathbb{N}, \Lambda^{\alpha}(\varnothing), \bar{\Lambda}^{\alpha}(\varnothing))$ where $\Lambda^{\alpha}(\varnothing)$ denotes the α-th iteration of the operator Λ, such that:

$$\Lambda^{0}(\varnothing) := \varnothing \qquad \Lambda^{\alpha+1}(\varnothing) := \Lambda(\Lambda^{\alpha}(\varnothing)) \qquad \Lambda^{\lambda}(\varnothing) := \bigcup_{\alpha<\lambda} \Lambda^{\alpha}(\varnothing)$$

for λ a limit ordinal. $\bar{\Lambda}^{\alpha}$ is defined similarly.

Λ and $\bar{\Lambda}$ are monotone operators in the same sense as Φ. By appealing to the Knaster-Tarski theorem, this means that there is a least S such that $\Lambda(S) = S$ [McG90, p. 107], and a least $\bar{S}$ such that $\bar{\Lambda}(\bar{S}) = \bar{S}$. We say S is a Kripke truth set iff it is a fixed point of Λ and $\bar{S}$ is a Kripke falsity set iff it is a fixed point of $\bar{\Lambda}$.

We define a *fixed point model* of this construction as $\mathfrak{K}^{\theta} := (\mathbb{N}, S, \bar{S})$ where S is a fixed point of Λ and $\bar{S}$ is a fixed point of $\bar{\Lambda}$. We can show that all the fixed points of $\mathfrak{K}^{\theta}$ coincide with the fixed points of $\mathfrak{K}_{\theta}$ (see [Hal14, p. 195-196])[5], by demonstrating that the Kripke truth set S is equivalent to the extension $\mathcal{E}$ in any fixed point, and, $\bar{S}$ is equivalent to the anti-extension $\mathcal{A}$. In other words:

Theorem 2.8. *$\Lambda^{\theta}(\varnothing) = S = \mathcal{E}$ and $\bar{\Lambda}^{\theta}(\varnothing) = \bar{S} = \mathcal{A}$ for $\theta \geqslant \omega_1^{CK}$*

Corollary 2.9. $\mathfrak{K}^{\omega_1^{CK}} = \mathfrak{K}_{\omega_1^{CK}} = \mathfrak{K}$

Thus the *least fixed point model* $\mathfrak{K}^{\omega_1^{CK}}$ that results from the positive inductive construction is equivalent to the model $\mathfrak{K}$ of Kripke's construction.

[5] In fact, this statement is a simplification, since the constructions are formulated in different languages; where Kripke's original construction only includes T, the alternative construction also includes F. This means the fixed points constructions are not identical. However, the falsity predicate F is definable in Kripke's construction as $F := T\neg$, and therefore the fixed points in the alternative construction can be directly translated into the original fixed point construction, and vice versa.

The benefit of this definition of truth, as opposed to Kripke's original approach, is that the iterative construction of the model is extended to the non-truth operators. Therefore, we can argue that for each logical connective, when a sentence containing that connective is added to the truth set, we move to a new stage in the construction. The stages in the construction of the Kripkean model in this sense are therefore different than the stages in Kripke's original construction, but the fixed point models are equivalent (see also [Mos74] and [McG90] for more on the properties of positive inductive definitions).

2.2 Axiomatic Truth Theories

The axiomatic approach offers an alternative to the semantic one. In these theories, we take truth as a primitive notion. An axiomatic truth theory is formed over some *base theory* by extending it with a truth predicate and a set of axioms for it. These axioms determine the behaviour of the truth predicate within the base theory. Therefore we avoid having to establish some metatheory in order to talk about the base theory, as in the semantic case. The usual (though not the only possible) choice of base theory is Peano arithmetic (PA), which is what all the theories in the remainder of this chapter will be formed over.

2.2.1 Disquotational theories

Despite Tarski's aversion to an axiomatic theory of truth, his inductive truth definitions have been used as a foundation for the axiomatic approach. One may attempt to axiomatize Tarski's definition by taking the approach initially rejected by himself: using the *Tarski biconditionals* arising from Convention T as the direct and only axioms of a theory. These biconditionals are also called *disquotation sentences* and are sentences of the form

$$A \leftrightarrow T\underset{\cdot}{A}$$

A theory based on these sentences is called a *disquotational theory*, and reflects the idea, which has been defended by philosophers such as Field in [Fie94], that truth is a device of disquotation and holds little further explanatory power.

Adding the truth biconditionals to the base theory PA, one obtains the theory TB (*Tarski biconditionals*), which is a typed theory, i.e., the application of the truth predicate is restricted to sentences of $\mathcal{L}$ [Hal14, p. 53].

Definition 2.10 (TB)**.** The theory TB consists of all axioms of PA_T, including all the induction axioms, as well as all sentences of the form $A \leftrightarrow T\underset{\cdot}{A}$, where A is a sentence of $\mathcal{L}$.

In TB, A must be a sentence containing no free variables. This restriction can be lifted to obtain the theory UTB (*uniform Tarski biconditionals*) allowing for *uniform disquotation sentences*.

Definition 2.11 (UTB)**.** The theory UTB consists of all axioms of PA_T, including all the induction axioms, as well as all sentences of the form $\forall x_1 ... \forall x_n (A(x_1^\circ, ..., x_n^\circ) \leftrightarrow T\underset{\cdot}{A}(x_1, ..., x_n))$, where $A(x_1, ... x_n)$ is a formula of $\mathcal{L}$ with $x_1, ..., x_n$ the free variables in A.

To transform TB and UTB into type-free theories, we cannot simply lift the restriction of T to truth-free sentences, since this would allow for the possibility of deriving the liar paradox [Hal14, p. 253]. One way to avoid this is to restrict the disquotation sentences to their *positive* form. This gives us the type-free theories PTB and PUTB.

The common criticism of disquotational theories, initially highlighted by Tarski, is that they are deductively too weak to form a sufficient theory. TB and UTB conservatively extend PA, meaning they prove all the same arithmetical statements as PA without proving any new content [Hal14, p. 55]. This might be in line with the disquotational notion of truth, but it does mean these theories are of little proof-theoretic value for saying anything novel about PA.

2.2.2 Compositional Truth

While Tarski rejected the idea of adding further truth axioms to get a stronger theory than the disquotational ones, this approach has become rather popular. Such theories are called *compositional axiomatic theories*, and they can result in much stronger truth axiomatizations than theories based merely on disquotation sentences. The most straightforward example of a typed compositional axiomatization is the theory CT (*Compositional Truth*), which directly translates the inductive clauses that can be obtained from Tarski's truth definition into axioms to be added over the base theory Peano arithmetic [Hal14, p. 65, p. 89].

Definition 2.12 (CT)**.** The theory CT consists of all axioms of PA_T, including all the induction axioms, as well as the following axioms formulated in $\mathcal{L}_T$[6]:

CT1 $\forall x. \forall y. (T(x \underset{\cdot}{=} y) \leftrightarrow x^\circ = y^\circ)$

CT2 $\forall x. (\mathrm{Sent}_{\mathcal{L}_T}(x) \rightarrow (T(\underset{\cdot}{\neg} x) \leftrightarrow \neg T x))$

CT3 $\forall x. \forall y. (\mathrm{Sent}_{\mathcal{L}_T}(x \underset{\cdot}{\wedge} y) \rightarrow T(x \underset{\cdot}{\wedge} y) \leftrightarrow Tx \wedge Ty)$

CT4 $\forall x. \forall y. (\mathrm{Sent}_{\mathcal{L}_T}(x \underset{\cdot}{\vee} y) \rightarrow T(x \underset{\cdot}{\vee} y) \leftrightarrow Tx \vee Ty)$

CT5 $\forall A. \forall x. (\mathrm{Sent}_{\mathcal{L}_T}(\underset{\cdot}{\forall} x. A) \rightarrow (T(\underset{\cdot}{\forall} x. A) \leftrightarrow \forall y. TA[y/x]))$

[6]Note that in CT1 (and in fact any of the axioms of CT), it is a requirement that x and y must be (Gödel codes of) closed terms in $\mathcal{L}_T$. This is apparent from the fact that we perform some operations over the Gödel codes of closed terms. However, rather than stating this requirement explicitly as part of the axiom, it is suppressed for simplicity's sake. We will continue to suppress this requirement for further definitions in this thesis. This is in line with several presentations of the axiomatic theories (cf. [Hal14]) and will also prove simpler for translating truth theories into proof systems for proof theoretic investigations in further chapters.

CT6 $\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}_T}(\underset{\cdot}{\exists} x.A) \to (T(\underset{\cdot}{\exists} x.A) \leftrightarrow \exists y.TA[y/x]))$

Restricting the occurrence of the truth predicate to typed sentences results in theories that are somewhat inexpressive and that do not give a complete description of truth in the full extent of a language. Therefore the development of type-free theories that allow the truth predicate to be applied to itself has been a particular area of interest. Type-free theories are more expressive and can give a more complete description of truth in a language. For the remainder of this chapter, we will shift our attention to the type-free case [7].

2.2.3 Kripke-Feferman

The theory *Kripke-Feferman* (KF), [Fef91], is an axiomatization of Kripke's semantic theory identifying truth as a (least) fixed point. Besides the truth predicate T, the presentation of KF we give here also uses a falsity predicate, F. Since F is definable by T, some authors choose to work with only the truth predicate T. However, since we believe having a truth and a falsity predicate more accurately translates the notion of an extension and anti-extension as in Kripke's construction, we choose to include F as its own predicate with corresponding rules.

A feature of KF is that T and F may only occur positively. Thus, we are formulating our theory in the language $\mathcal{L}^+_{T,F}$. We let $\mathrm{PA}_{T,F}$ be Peano arithmetic extended with T and F.

Definition 2.13 (KF)**.** [8] The theory KF consists of all axioms of $\mathrm{PA}_{T,F}$ including induction, as well as the following axioms formulated in $\mathcal{L}^+_{T,F}$:

KF1 $\forall x.\forall y.(((T(x \underset{\cdot}{=} y) \leftrightarrow x^\circ = y^\circ) \land (T(\underset{\cdot}{\neg}(x \underset{\cdot}{=} y)) \leftrightarrow \neg x^\circ = y^\circ)))$

KF2 $\forall x.\forall y.(((F(x \underset{\cdot}{=} y) \leftrightarrow \neg x^\circ = y^\circ)) \land (F(\underset{\cdot}{\neg}(x \underset{\cdot}{=} y)) \leftrightarrow x^\circ = y^\circ))$

KF3 $\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\land} y) \to (T(x \underset{\cdot}{\circ} y) \leftrightarrow Tx \circ Ty))$ for $\circ \in \{\land, \lor\}$

KF4 $\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\land} y) \to (F(x \underset{\cdot}{\circ} y) \leftrightarrow Fx \bar{\circ} Fy))$ for $\circ \in \{\land, \lor\}$

KF5 $\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} x.A) \to (T\underset{\cdot}{\mathsf{Q}} x.A \leftrightarrow \mathsf{Q} y.TA[y/x]))$ for $\mathsf{Q} \in \{\forall, \exists\}$

KF6 $\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} x.A) \to (F\underset{\cdot}{\mathsf{Q}} x.A \leftrightarrow \bar{\mathsf{Q}} y.FA[y/x]))$ for $\mathsf{Q} \in \{\forall, \exists\}$

KF7 $\forall x.((T\underset{\cdot}{T} x \leftrightarrow Tx) \land (T\underset{\cdot}{F} x \leftrightarrow Fx))$

---

[7] For more information on CT and other typed theories, cf. Halbach [Hal14].

[8] The definition for KF used here is adapted from [Lei16, p. 575]. Note that the axiom KF9 is not always included in KF. For our purposes it will prove useful to assume this as part of KF, hence the inclusion. Furthermore, as in the definition of CT, we suppress the requirement that x and y must be closed terms of $\mathcal{L}^+_{T,F}$ for each axiom, but specifically KF1, KF2 and KF7, KF8. This requirement is left implicit.

KF8 $\forall x.((F\dot{T}x \leftrightarrow Fx) \wedge (F\dot{F}x \leftrightarrow Tx))$

KF9 $\forall x.((Tx \vee Fx) \rightarrow \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x))$

where $\bar{\wedge} = \vee$, $\bar{\vee} = \wedge$, $\bar{\forall} = \exists$ and $\bar{\exists} = \forall$.

The axioms KF7 and KF8 of this theory capture the construction of Kripke's extension and anti-extension in the partial model and make it a type-free theory. KF9 says that every term that is true or false must be a sentence of $\mathcal{L}^+_{T,F}$, so only sentences can have a truth value.

In order to force KF's consistency, we could add an additional axiom:

$$\textsc{Cons}\ \forall x \neg(Tx \wedge Fx)$$

Cons makes it impossible for KF to contain truth value gluts, which is in line with Kripke's original construction that does not allow these gluts. It was shown by Cantini [Can89] that including the consistency axiom in KF does not affect its proof-theoretic strength, therefore proof-theoretically KF+Cons is equivalent to KF. Some authors include this axiom in their formulation of KF. However, we have chosen to omit it, aligning with Feferman's original formulation of the theory in [Fef91].

KF is roughly the result of adapting the clauses of the Kripke truth-falsity pair $(S, \bar{S})$ as defined in Definition 2.7 into axioms. In fact, the classical model $(\mathbb{N}, S, \bar{S})$ is a model of KF iff $(S, \bar{S})$ is a Kripke truth-falsity pair [Hal14, p. 197]. By definition, if S is a fixed point of Λ and $\bar{S}$ of $\bar{\Lambda}$, then S and $\bar{S}$ verify all the axioms of KF. In the other direction, if $(\mathbb{N}, S, \bar{S})$ is a model of KF then it can be verified by the axioms of KF that the extension of T is a Kripke truth set, likewise the anti-extension of F is a Kripke falsity set.

Thus, despite being based on the partial model obtained through the strong Kleene scheme, KF is formulated in classical logic. We can call the inner logic of KF the set of sentences A such that $T\underset{\cdot}{A}$ is a theorem of the system, and its outer logic the set of sentences provable in the system. These logics do not match; the outer logic of KF proves all classical sentences, including the tautology $L \vee \neg L$, where L is the liar sentence, but this statement is not provable in its inner logic, which is partial and leaves L undecided [Hal14, p. 215]. Thus KF proves theorems which are untrue according to the notion of truth as represented in Kripke's construction. In fact, this means KF is only sound with respect to Kripke's model in the following sense:

Theorem 2.14. *([HH06, p. 682]) If* KF $\vdash T\underset{\cdot}{A}$*, then A holds in all Kripke fixed point models.*

It is not possible to force the inner and outer logic of KF to coincide, as doing so would require it to be closed under the following rules:

$$\textsc{Nec}\quad \frac{A}{T\underset{\cdot}{A}} \qquad\qquad \textsc{Conec}\quad \frac{T\underset{\cdot}{A}}{A}$$

By adding Nec and Conec to a theory we can assert that if a sentence is provable in a theory, it is also true in the theory, and vice versa. However, Halbach and Horsten have shown that the result of closing KF under both Nec and Conec is inconsistent [HH06, p. 202]. This discrepancy between the inner and outer logic of KF makes KF somewhat questionable as a type-free truth theory, since T and F in KF do not accurately reflect the semantic notions of truth and falsity they ought to describe.

Despite this, KF is still a valuable theory for describing the reflective closure of PA, that is, for making explicit any implicit assumptions that come with the commitment to PA (cf. [Hal14, p. 312-315] for more). Furthermore, it is a practically useful theory. Deductively KF is stronger than the disquotational theories[9] and the typed theory CT, as well as its other type-free counterpart the theory FS[10]; in fact KF is proof-theoretically equivalent to the theory of ramified truth up to the level ε_0 according to Feferman [Hal14, p. 209]. Furthermore, KF is sound with respect to arithmetic [Hor11, p. 126].

2.2.4 Systems related to KF

Partial axiomatizations of Kripke's theory: IKF and PKF

Several authors have argued for adjusted versions of KF which more closely adhere to Kripke's partial notion of truth. Among them is Reinhardt, who proposes the theory IKF (for *inner* KF), where IKF $= \{A : \text{KF} \vdash T\underset{\cdot}{A}\}$ [Rei86, p. 232]. This theory is arithmetically equivalent to KF and it does not contain the liar sentence [Hor11, p. 129]. However, it is also not really a proper theory in the sense that it does not directly axiomatize a notion of truth; rather it is just a computably enumerable set of sentences based on the theory KF.

Halbach and Horsten [HH06] propose a different axiomatization of Kripke's theory formulated in partial logic, called PKF (for *partial* KF). Their system is given in the style of a sequent calculus, where the truth principles of PKF are given as rules that reflect the truth axioms of KF, but formulated in partial logic. Halbach and Horsten argue that PKF is the most natural axiomatization of Kripke's theory [HH06, p. 679]. Certainly, since PKF builds its rules on top of a system of partial logic, its inner and outer logic actually coincide, unlike KF, and therefore the system better reflects Kripke's partial truth models. PKF is a subtheory of KF, thus arithmetically sound and it leaves the truth value of the liar paradox undecided, meaning it is sound with respect to Kripke's fixed point model. However, this system is proof-theoretically much weaker than KF [HH06, p. 701]. Furthermore, we can argue that it is more difficult to work with than KF, since reasoning in partial logic feels less natural than classical reasoning, making it somewhat unattractive as a practical

[9] PTB and PUTB are subtheories of KF [Can89, p. 102]. However, it can be shown that PUTB proves all the same arithmetical sentences as KF, thus implying its truth-free content is equivalent [Hal14, p. 268] (cf. [Hal01] for more on the relation between disquotational and compositional theories.)

[10] The theory Friedman-Sheard generalizes the axioms of CT to their type-free counterparts. In doing so it retains the classical axiom $\forall t(T\neg t \leftrightarrow \neg Tt)$ which KF rejects, but in turn loses the axiom $\forall A(TT\underset{\cdot}{t} \leftrightarrow Tt)$. FS is not further discussed here (cf. Halbach [Hal14] for further information).

theory.

An axiomatization of Kripke's theory for the least fixed point: μKF

Another criticism of KF as an axiomatization of Kripke's semantics is that its axioms describe all the fixed points of Kripke's partial model, and not just the minimal one. This is problematic, since in Kripke's model truth is defined by the least fixed point. An axiomatization that can only capture any arbitrary fixed point will not suffice to describe truth as originally intended. For this reason, Burgess [Bur14] attempts to formulate a version of KF that more directly incorporates minimality. This system, called μKF, takes only the right-to-left direction of KF's truth axioms KF1-KF8 as in Definition 2.13 (the compositional direction). Thus, for example it is an axiom of μKF that if x is true and y is true, then their conjunction is true, but the other direction does not necessarily hold [Hal14, p. 245]. It then further extends KF with an axiom scheme for minimality, stating that for every formula $P(x)$ of $\mathcal{L}^{+}_{T,F}$, if the set of true sentences satisfying P is closed under the compositional truth axioms KF1-KF8, then every true sentence satisfies P [Bur14, p. 19] (and similarly for false sentences). To formally define this axiom scheme, we must first construct the sentence $M(P,Q)$, which gives the conjunction of all compositional axioms of KF with T replaced by a formula $P(x)$, and F by $Q(x)$, except for in the last clause.

Definition 2.15. ([Hal14, p. 244]) We define M as: $M(P,Q) :=$

$$\forall x.\forall y.(x^\circ = y^\circ \to P(x \underset{\cdot}{=} y))\wedge$$

$$\forall x.\forall y.(\neg x^\circ = y^\circ \to Q(x \underset{\cdot}{=} y))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(x \underset{\cdot}{\wedge} y) \to (P(x) \wedge P(y) \to P(x \underset{\cdot}{\wedge} y)))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(x \underset{\cdot}{\wedge} y) \to (Q(x) \vee Q(y) \to Q(x \underset{\cdot}{\wedge} y)))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(x \underset{\cdot}{\vee} y) \to (P(x) \vee P(y) \to P(x \underset{\cdot}{\vee} y)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(x \underset{\cdot}{\vee} y) \to (Q(x) \wedge Q(y) \to Q(x \underset{\cdot}{\vee} y)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(\underset{\cdot}{\forall} x.A) \to (\forall y.P(A[y/x]) \to P(\underset{\cdot}{\forall} x.A)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(\underset{\cdot}{\forall} A.x) \to (\exists y.Q(A[y/x]) \to Q(\underset{\cdot}{\forall} x.A)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(\underset{\cdot}{\exists} A.x) \to (\exists y P(A[y/x]) \to P(\underset{\cdot}{\exists} A.x)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^{+}_{T,F}}(\underset{\cdot}{\exists} A.x) \to (\forall y Q(A[y/x]) \to Q(\underset{\cdot}{\exists} y x)))\wedge$$

$$\forall t.((P(t^\circ) \to P(\underset{\cdot}{T} t)) \wedge \forall t.(Q(t^\circ) \to P(\underset{\cdot}{F} t)))\wedge$$

$$\forall t.((Q(t^\circ) \rightarrow Q(\underset{\cdot}{T}t)) \land \forall t.(P(t^\circ) \rightarrow Q(\underset{\cdot}{F}t)))$$

Definition 2.16 (μKF)**.** The theory μKF consists of all axioms of KF, where for KF1-KF8 we only take the compositional direction, as well as the axiom:

> MIN $M(P,Q) \rightarrow \forall x.((Tx \rightarrow P(x)) \land (Fx \rightarrow Q(x))$ for all formulae $P(x), Q(x)$ of $\mathcal{L}^+_{T,F}$.

From the minimality scheme it is possible to derive the decompositional direction of the axioms of KF [Bur14, p. 19]. μKF is proof-theoretically a stronger theory than KF, since it contains KF but also derives theorems that KF does not contain. For instance, μKF derives the consistency axiom CONS, which cannot be done in KF [Bur14, p. 19]. From this it follows that μKF can prove the liar sentence, which follows from the consistency axiom [Hal14, p. 247].[11] In fact, Burgess shows that μKF is proof-theoretically as strong as the theory ID_1, which is a theory stating the existence of a fixed point for each positive arithmetic operator [Bur14, p. 21]. μKF does, however, contain the same issue as KF, in that its inner logic is partial but its outer logic is classical.

An axiomatizion of Kripke's theory based on a supervaluation scheme: VF

Kripke's theory uses the strong Kleene evaluation scheme to decide the truth value of sentences. However, it is possible to formulate the theory using an alternative evaluation scheme, providing different criteria for validity. One such alternative scheme is the *supervaluation* scheme. The idea behind supervaluations is that we can obtain classical variant $\mathfrak{C}$ of our partial model $\mathfrak{K}$, by adding sentences whose truth value is undetermined in $\mathfrak{K}$ to the extension [Hal14, p. 250]. According to the supervaluation scheme $\mathfrak{K} \models_{sv} A$ if $\mathfrak{C} \models_{sv} A$. Thus, through supervaluations we get rid of the truth gaps in our model and obtain what is essentially a classical model, which avoids the degree of ambiguity about the truth of some statements in our partial truth definition.

Cantini [Can90] has axiomatized Kripke's theory using supervaluation schemes into the theory VF. This theory proves all classical tautologies. We can, for instance, determine that the classical tautology $T(L \underset{\cdot}{\lor} \neg L)$, where L is the liar sentence, is a theorem of the outer logic of VF [HL22]. Cantini showed that his theory, like μKF, is proof-theoretically equivalent to ID_1 [Can90, p. 258]. However, despite what might be expected given the supervaluation construction of VF, the inner and outer logic of VF are actually incompatible, demonstrated by the fact that it cannot be consistently closed under NEC and CONEC [Hal14, p. 251]. VF, therefore, suffers from some of the same flaws as KF.

---

[11] The liar sentence that is proved in μKF is in the sense that $\neg\lambda$ is in the Kripke model iff it is in the extension of T. Given the restrictions to our language, the fact that the liar can be proven does not immediately lead to the theory being inconsistent.

Chapter 3

Cyclic Proof Theory

3.1 Foundations of Cyclic Proof Theory

In proof theory we often have to rely on inductive reasoning to construct infinitary arguments. This makes inductive principles powerful tools that significantly increase the proof-theoretic strength of our system. The most straightforward (and by far the most common) way to formally implement induction into a finite proof system is to explicitly add some kind of induction rule [Bro05, p. 78], resulting in what we will call *inductive* proof systems that can finitely represent possibly infinite arguments. However, induction rules, unlike the logical rules, cannot be applied mechanically. Implementing them requires some additional creativity and decision on the part of the prover, by, for example, guessing the right inductive hypothesis. Such tasks are especially challenging in the setting of automated reasoning, as the approach is not always straightforward. Therefore automated provers, for instance, often fail to cope with inductive reasoning, limiting the complexity of problems that they can solve [BGP12, p. 350].

Therefore, in recent years and as we want to improve systems for automated reasoning, it has become relevant to find alternative approaches for expressing inductive reasoning in formal proofs that avoid these difficulties. We might, for example, circumvent the use of induction rules by instead allowing for *ill-founded* reasoning. Ill-founded proof systems introduce the possibility of giving proofs by *infinite descent*, in other words proofs that may have branches that expand infinitely upwards [Sim17, p. 287]. Proofs by infinite descent have been shown to be proof-theoretically equivalent to inductive proofs in certain settings (cf. [Sim17]). A particular kind of ill-founded proofs are *cyclic proofs*, which use circular reasoning to describe derivations that may be infinite. Thus cyclic proofs are represented as finite graphs that may include cycles, usually over some specified fixed point of the theory. If we were to unfold these cycles, the graph would be transformed into an infinitely descending proof tree with finitely many distinct subproofs; we call this a *regular* ill-founded proof tree. Thus cyclic proofs describe regular ill-founded proofs [Sim17, p.

291]. Cyclic proofs can simulate inductive reasoning, and thereby avoid difficulties that come with the use of explicit induction rules.

For a simple example of a cyclic proof, we consider the irrationality of $\sqrt{2}$. Informally, we can prove that $\sqrt{2}$ is not a rational number through an argument of infinite descent. This can be formalized into a cyclic sequent calculus derivation, where we introduce a rule Subst that substitutes one variable in the sequent for another, which allows for us to get identical nodes over which we can cycle. This derivation is given below, with some steps omitted for simplicity and with $(\star)$ denoting the presence of a cycle between two nodes:

$$
\begin{prooftree}
\AxiomC{$0 < x, x^2 = 2y^2 \Rightarrow \bot(\star)$}
\LeftLabel{\textsc{Subst}}
\UnaryInfC{$0 < y, y^2 = 4z^2 \Rightarrow \bot$}
\UnaryInfC{$y < x, 0 < y, 4z^2 = 2y^2 \Rightarrow \bot$}
\AxiomC{}
\LeftLabel{\textsc{Ax}}
\UnaryInfC{$0 < x, x^2 = 2y^2 \Rightarrow 0 < y < x \land \exists z.(x = 2z)$}
\LeftLabel{\textsc{Cut}}
\BinaryInfC{$0 < x, x^2 = 2y^2 \Rightarrow \bot(\star)$}
\end{prooftree}
$$

Figure 3.1: A cyclic proof of the irrationality of $\sqrt{2}$ [Sim17, p. 287]

More formally, cyclic proofs can be seen as finite representations of ill-founded trees with finite branching and additional back-edges mapping leaves to inner nodes; where this occurs we call the leaf a *bud* and the inner node it maps to its *companion*. Formally, we say a *cyclic tree* is a pair $C = (P, \beta)$ consisting of a finite proof tree P and a partial function $\beta : \mathit{Leaf}(P) \to \mathit{Inner}(P)$ mapping a leaf node to an inner node of P.

In general, the idea of building proofs on the basis of circular argumentation is cause for suspicion, since circular reasoning can easily lead to fallacies and inconsistencies. Therefore, constraints need to be put on a cyclic proof to make sure it is actually expressing a sound argument; it is usually not sufficient to simply rely on the soundness of the derivation rules. Take for instance the following cyclic derivation, where $(\star)$ denotes the presence of a cycle between two nodes:

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{\textsc{Ax}}
\UnaryInfC{$\bot \Rightarrow \bot$}
\AxiomC{$\Rightarrow \bot(\star)$}
\RightLabel{\textsc{Cut}}
\BinaryInfC{$\Rightarrow \bot(\star)$}
\end{prooftree}
$$

Figure 3.2: An inconsistent cyclic derivation

Without any further restrictions being placed on the proof, the argument would be considered sound; and yet we derive a contradiction [Weh23, p. 6]. We therefore need to make sure our proof does not result in a logical fallacy. A standard approach do this and to ensure soundness is to define a global condition on cyclic proofs, a so-called *progressivity condition*. The specific progressivity conditions for cyclic and ill-founded proofs are adapted to the system being considered, but will generally implement the same idea of making sure our proof does not end up in a vicious cycle like in Figure 3.2. Below we will illustrate this for the system of cyclic Peano arithmetic.

Some of the fundamental results in cyclic proof theory are given by Brotherston [Bro05], who investigates a cyclic formulation of inductive first-order logic, and Simpson [Sim17], who expands on this by giving a cyclic proof system of Peano arithmetic (CPA) that is equivalent to Peano arithmetic. In recent years, these results have been expanded upon considerably and cyclic proof theory has evolved into an active field of study within the larger realm of proof theory. It has also established itself as a valuable tool in practical applications for, for instance, automatic theorem proving (cf. [BGP12], [TB20], [TU22], [BDP11] for some applications). Additionally, cyclic proofs have played a part in proof-theoretic investigations into inductive systems, like modal μ-calculus (cf. [AL16], [AEL22]), arithmetical inductive definitions (cf. [DM23]) or temporal logics (cf. [KS16], [ALMT23]).

3.2 Cyclic Peano arithmetic

As mentioned above, the cyclic system CPA was initially developed by Simpson in [Sim17], where he also showed this system to be sound and equivalent to the inductive proof system PA. The descriptions of these systems in this section are primarily taken from Wehr's presentation of Simpson's system CPA in [Weh23][1].

3.2.1 The system Peano arithmetic

In order to define the cyclic system CPA, we first define an inductive proof system for Peano arithmetic.

Definition 3.1 (Peano arithmetic)**.** A *sequent* $\Gamma \Rightarrow \Delta$ in Peano arithmetic is a pair of finite sets Γ and Δ of first-order formulae. Let A, B be formulae of first-order arithmetic; and s, t terms of first-order arithmetic. The proof system *Peano arithmetic* (PA) consists of the following derivation rules and axioms:

Logical rules:

$$\mathrm{L}_{\neg}\ \frac{\Gamma \Rightarrow \Delta, A}{\Gamma, \neg A \Rightarrow \Delta} \qquad\qquad \mathrm{R}_{\neg}\ \frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg A}$$

$$\mathrm{L}_{\wedge}\ \frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta} \qquad\qquad \mathrm{R}_{\wedge}\ \frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \wedge B}$$

[1] Simpson's version of CPA uses slightly different cyclic rules than the ones that will be presented here. As a result, Simpson's soundness proof, as well as his derivation of the PA-induction axiom in CPA, also look different, though the outline of the argument is the same.

$$\mathrm{L}_{\vee}\ \frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta} \qquad \mathrm{R}_{\vee}\ \frac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}$$

$$\mathrm{L}_{\rightarrow}\ \frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma, B \Rightarrow \Delta}{\Gamma, A \rightarrow B \Rightarrow \Delta} \qquad \mathrm{R}_{\rightarrow}\ \frac{\Gamma, A \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \rightarrow B}$$

$$\mathrm{L}_{\forall}\ \frac{\Gamma, A[t/x] \Rightarrow \Delta}{\Gamma, \forall x.A \Rightarrow \Delta} \qquad \mathrm{R}_{\forall}\ \frac{\Gamma \Rightarrow \Delta, A[y/x]}{\Gamma \Rightarrow \Delta, \forall x.A}\ y \notin \mathrm{FV}(\Gamma, \Delta, \forall x.A)$$

$$\mathrm{L}_{\exists}\ \frac{\Gamma, A[y/x] \Rightarrow \Delta}{\Gamma, \exists x.A \Rightarrow \Delta}\ y \notin \mathrm{FV}(\Gamma, \Delta, \exists x.A) \qquad \mathrm{R}_{\exists}\ \frac{\Gamma \Rightarrow \Delta, A[t/x]}{\Gamma \Rightarrow \Delta, \exists x.A}$$

$$\mathrm{L}_{=}\ \frac{\Gamma[t,s] \Rightarrow \Delta[t,s]}{\Gamma[s,t], s = t \Rightarrow \Delta[s,t]} \qquad \mathrm{R}_{=}\ \frac{}{\Gamma \Rightarrow t = t, \Delta}$$

$$\mathrm{L}_{\perp}\ \frac{}{\Gamma, \perp \Rightarrow \Delta}$$

Structural rules:

$$\textsc{Wk}\ \frac{\Gamma \Rightarrow \Delta}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'} \qquad \textsc{Cut}\ \frac{\Gamma, A \Rightarrow \Delta \qquad \Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta}$$

Axioms for first-order arithmetic:

$$\begin{array}{rr} \Rightarrow 0 \neq s(t) & \Rightarrow t + s(s) = s(t + s) \\ s(t) = s(s) \Rightarrow t = s & \Rightarrow t \cdot 0 = 0 \\ \Rightarrow t + 0 = t & \Rightarrow t \cdot s(s) = (t \cdot s) + t \end{array}$$

Induction axiom:

$$A(0), \forall x.(A(x) \rightarrow A(s(x))) \Rightarrow A(t)$$

3.2.2 The system CPA

We now introduce a sequent calculus for the cyclic system CPA. This system consists primarily of the same rules as the inductive sequent calculus formulation of PA, but replaces its induction axiom with some rules imposing circular reasoning.

Definition 3.2 (Cyclic Peano arithmetic)**.** The proof system *cyclic Peano arithmetic* (CPA) consists of all the sequents, rules and axioms of PA, except for the induction axiom, and is extended with the following rules [Weh23, p. 5]:

$$\textsc{Subst}\ \frac{\Gamma \Rightarrow \Delta}{\Gamma[t/x] \Rightarrow \Delta[t/x]}\ x \notin \mathrm{FV}(t)$$

$$\textsc{Case}_x\ \frac{\Gamma[0/x] \Rightarrow \Delta[0/x] \qquad \Gamma[s(x)/x] \Rightarrow \Delta[s(x)/x]}{\Gamma \Rightarrow \Delta}\ x \notin \mathrm{FV}(\Gamma, \Delta)$$

Definition 3.3 (CPA pre-proof, [Weh23])**.** A *pre-proof* in CPA is a pair $\Pi = (C, \lambda)$ where C is a cyclic tree, and λ is a function $\lambda : Node(P) \to Seq$ labelling the nodes of P with sequents, such that each inner node is associated with one of the CPA derivation rules, leaves are labelled by axioms and buds share a label with its companion.

As previously demonstrated by the fallacy in Figure 3.2, this definition does not suffice to construct a sound system. Hence, in the definition of a CPA proof we introduce a global progressivity condition to achieve soundness. Though a progressivity condition for CPA could be formulated in several ways, we consider here a progressivity condition as formulated in [Weh23, p. 6].

Definition 3.4 (Infinite branch)**.** Let Π be a CPA pre-proof. An *infinite branch* through Π is an infinite sequence of nodes $(t_i)_{i \in \omega}$ of the proof tree P, where for each $i \geqslant 0$, either t_{i+1} is a node of some premise of t_i, or t_i is a bud node and t_{i+1} is its companion. This induces an infinite sequence of sequents $(\Gamma_i \Rightarrow \Delta_i)_{i \in \omega}$.

Definition 3.5 (Progressing arithmetic trace)**.** A variable x has an *arithmetic trace* along an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i \in \omega}$ if there exists an $n \in \omega$ such that x is a free variable of Γ_i and Δ_i for all $i \geqslant n$. This trace is *progressing* if it passes through a $\textsc{Case}_x$-rule infinitely often for the same variable x.

Definition 3.6 (CPA proof)**.** A pre-proof in CPA Π is a *proof* in CPA if for every infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i \in \omega}$ there exists some variable which has a progressing trace along it.

Asserting the progressivity condition for a cyclic proof allows us to avoid circular reasoning leading to contradictions as in Figure 3.2. In this example, we have one infinite branch, but the $\textsc{Case}_x$ rule does not occur on that branch for any variable, therefore any trace along this branch is not progressing. So that pre-proof does not qualify as a proof in CPA.

3.2.3 Soundness of CPA

We demonstrate that the proof system CPA is sound with regard to $\mathbb{N}$. To this purpose, we will first prove the following lemma:

Lemma 3.7 (Reflecting falsity)**.** *Given an inference step*

$$\mathsf{r}\ \frac{\Gamma_0 \Rightarrow \Delta_0 \qquad \cdots \qquad \Gamma_n \Rightarrow \Delta_n}{\Gamma \Rightarrow \Delta}$$

If there exists an environment $\rho : FV(\Gamma) \cup FV(\Delta) \to \omega$ *interpreting the free variables of any formula* $A \in \Gamma \cup \Delta$ *as natural numbers, such that* $\rho \models \bigwedge \Gamma$ *and* $\rho \not\models \bigvee \Delta$ *then there is* $\rho' \leqslant \rho$ *and* $0 \leqslant i \leqslant n$ *such that* $\rho' \models \bigwedge \Gamma_i$ *and* $\rho' \not\models \bigvee \Delta_i$*.*

Proof. We show that for any conclusion of a rule $\Gamma \Rightarrow \Delta$, if $\rho \not\models \Gamma \Rightarrow \Delta$, there is some environment ρ' such that $\rho' \not\models \Gamma_i \Rightarrow \Delta_i$, where $\Gamma_i \Rightarrow \Delta_i$ is some premise of the rule. We look at some illustrative cases[2].

- $\mathrm{R}_\wedge$: Then $\Delta = \Delta_0, A \wedge B$ and the premises are $\Gamma \Rightarrow \Delta_0, A$ and $\Gamma \Rightarrow \Delta_0, B$. We can fix $\rho' := \rho$. If $\rho \not\models \Gamma \Rightarrow \Delta$ then any $C \in \Delta$ is such that $\rho \not\models C$. In particular, $\rho \not\models A \wedge B$, so either $\rho \not\models A$ or $\rho \not\models B$. Then, either $\rho \not\models \Gamma \Rightarrow \Delta_0, A$ or $\rho \not\models \Gamma \Rightarrow \Delta_0, B$.

- $\mathrm{R}_\forall$: Then $\Delta = \Delta_0, \forall x.A$ and the premise is $\Gamma \Rightarrow \Delta_0, A[y/x]$ where y is a fresh variable, not occurring in Γ or Δ. If $\rho \not\models \Gamma \Rightarrow \Delta$ then $\rho \not\models \forall x.A$, so there exists an $n \in \omega$ such that, fixing $\rho' := \rho[y \mapsto n]$, we have $\rho' \not\models A[y/x]$. Then we have $\rho' \not\models \Gamma_0 \Rightarrow \Delta_0, A[y/x]$.

- $\mathrm{L}_=$: Then $\Gamma = \Gamma_0[s/x, t/y], s = t$, and $\Delta = \Delta_0[s/x, t/y]$. The premise is $\Gamma_0[t/x, s/y] \Rightarrow \Delta_0[t/x, s/y]$. Since $\rho \not\models \Gamma[s/x, t/y], s = t \Rightarrow \Delta[s/x, t/y]$ we know that $\rho \models s = t$, so for any term u, $\rho \models u[s/x] = u[t/x]$ and $\rho \models u[t/y] = u[s/y]$. So $\rho \not\models \Gamma_0[t/x, s/y] \Rightarrow \Delta_0[t/x, s/y]$.

- $\textsc{Subst}$: Then $\Gamma = \Gamma_0[t/x]$ and $\Delta = \Delta_0[t/x]$. The premise is $\Gamma_0 \Rightarrow \Delta_0$. We can fix $\rho' := \rho[x \mapsto t]$. Then for any variable $y \neq x$, $\rho' := \rho$, and for x we have $\rho' \models x$ iff $\rho \models t$. Therefore, for any formula A, $\rho \models A[t/x]$ iff $\rho' \models A(x)$. So, since $\rho \not\models \Gamma_0[t/x] \Rightarrow \Delta_0[t/x]$ then $\rho' \not\models \Gamma_0 \Rightarrow \Delta_0$.

- $\textsc{Case}_x$: x is not free in $\Gamma \cup \Delta$ and $\rho \not\models \Gamma \Rightarrow \Delta$. By $\textsc{Case}_x$ we have either $\rho(x) = 0$ or $\rho(x) = s(n)$ for some $n \in \omega$. If $\rho(x) = 0$ then we fix $\rho' := \rho$ so that $\rho \not\models \Gamma[0/x] \Rightarrow \Delta[0/x]$. If $\rho(x) = s(n)$ then we can fix $\rho' := \rho[x \mapsto n]$. Then $\rho'(x) = n < s(n)$. It follows that, since $\rho \not\models \Gamma(x) \Rightarrow \Delta(x)$, then $\rho' \not\models \Gamma[s(x)/x] \Rightarrow \Delta[s(x)/x]$.

The remaining cases are similar or trivial. In particular, in any other case we can fix $\rho' := \rho$ as in the first case, thus $\rho \not\models \Gamma_i \Rightarrow \Delta_i$ for some premise of any of these rules. □

Using Lemma 3.7, we can show that a CPA proof is sound, i.e. that when a statement can be proven in CPA, this statement is provably true in the standard model of first-order arithmetic $\mathbb{N}$.

Theorem 3.8 (Soundness)**.** [3] *If* CPA $\vdash \Gamma \Rightarrow \Delta$ *then* $\mathbb{N} \models \Gamma \Rightarrow \Delta$.

---

[2] These cases are also given in [Weh23, p. 7-8]

[3] See also [Sim17, p. 290] for Simpson's original formulation of the proof of soundness for CPA, and [Weh23, p. 7-8] for a presentation closer to the one given here.

Proof. Let Π be a CPA proof of $\Gamma \Rightarrow \Delta$ and suppose towards contradiction that $\mathbb{N} \not\models \Gamma \Rightarrow \Delta$. By iteration of Lemma 3.7, we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ through the proof, and a family of environments $(\rho_i)_{i\in\omega}$ such that $\rho_i : FV(\Gamma_i) \cup FV(\Delta_i) \to \omega$, and $\rho_{i+1} \leqslant \rho_i$. By the progressivity condition of Π we know that the infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ has a progressing arithmetic trace. This means that, by Definition 3.5, the trace is infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x. By the proof of Lemma 3.7, whenever $\rho_i \not\models \Gamma \Rightarrow \Delta$ and we pass through the righthand side of $\textsc{Case}_x$, we have $\rho_{i+1} \not\models \Gamma \Rightarrow \Delta$ where $\rho_{i+1} < \rho_i$. Therefore since we pass through this rule infinitely often, we have $\rho_{i+1}(x) < \rho_i(x)$ for infinitely many i, yielding an infinitely decreasing sequence that contradicts the well-foundedness of the natural numbers. Thus, whenever there is a proof Π of a sequent $\Gamma \Rightarrow \Delta$, it must be the case that $\mathbb{N} \models \Gamma \Rightarrow \Delta$. □

3.2.4 Equivalence of CPA and PA

In addition to soundness, we can also prove that any proof in CPA can be transformed into a proof in the inductive system PA, and vice versa. In other words, CPA is equivalent to PA:

Theorem 3.9 ([Sim17])**.** CPA = PA

This result is originally due to Simpson [Sim17]. The only difference between the rules of PA and CPA is that CPA uses the cyclic rules $\textsc{Subst}$ and $\textsc{Case}_x$ to make inductive arguments, whereas PA relies on the presence of the induction axiom.. Therefore, proving equivalence of the two systems requires us to prove these rules somehow commute with each other.

We give here the proof of one direction of the equivalence but omit the proof for the other direction, which is more involved. For the full proof of this direction, see the original paper by Simpson [Sim17].

Theorem 3.10. PA $\subseteq$ CPA.

Proof. Since the only difference between a PA proof and a CPA proof is that the former may include leaves labelled by the induction axiom, and the latter contains the cyclic rules $\textsc{Case}_x$ and $\textsc{Subst}$ instead, if we show that the induction axiom can be derived by these the CPA rules, then any PA proof can be simulated by a CPA proof. The induction axiom is derived in CPA as follows, with Γ denoting an abbreviation of $A(0), \forall x.(A(x) \to A(s(x))$:

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$\Gamma \Rightarrow A(0)$}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$\Gamma, A(s(x)) \Rightarrow A(s(x))$}
\AxiomC{$\Gamma \Rightarrow A(x)(\star)$}
\LeftLabel{Wk}
\UnaryInfC{$\Gamma \Rightarrow A(s(x)), A(x)$}
\LeftLabel{$\mathrm{L}_{\rightarrow}$}
\BinaryInfC{$\Gamma, A(x) \rightarrow A(s(x)) \Rightarrow A(s(x))$}
\LeftLabel{$\mathrm{L}_{\forall}$}
\UnaryInfC{$\Gamma \Rightarrow A(s(x))$}
\LeftLabel{$\textsc{Case}_x$}
\BinaryInfC{$\Gamma \Rightarrow A(x)(\star)$}
\LeftLabel{$\textsc{Subst}$}
\UnaryInfC{$\Gamma \Rightarrow A(t)$}
\end{prooftree}
$$

Figure 3.3: Derivation of the PA-induction axiom in CPA [Weh23, p. 6]

The nodes marked with $(\star)$ decribe the leaf node is a bud and the inner node is its companion. Now, any occurrence of the induction axiom in a proof of PA can be replaced by the above derivation. Since in the resulting derivation we pass through the $\textsc{Case}_x$ rule infinitely often, there is a progressing trace as in Definition 3.5 and therefore by Definition 3.6 the derivation is a CPA proof. $\square$

Chapter 4

Proof-Theoretic Representations of KF

We will now begin to turn our attention to the development of a cyclic proof system for the truth theory KF. In order to do so, in this chapter we first develop an inductive proof system for KF in the form of a finitary sequent calculus.[1] We additionally give inductive systems for the theory of truth as a Kripke minimal fixed point, μKF, and develop a proof system for a new theory that takes truth as a greatest fixed point of the Kripke construction, which we shall call νKF.

4.1 The system KF

Since in the theory KF we take Peano arithmetic as a base theory, in a proof-theoretic representation of KF we retain the rules of PA, on top of which we add a set of truth and falsity rules and axioms. Thus the finitary sequent calculus for KF consists of all the logical, arithmetic and structural sequent calculus rules and axioms of the inductive system PA, extended to the language $\mathcal{L}_{T,F}$, as well as a set of compositional truth and falsity rules that describe the axioms of KF as in 2.13.

Definition 4.1 (KF)**.** A *sequent* $\Gamma \Rightarrow \Delta$ in the Kripke-Feferman proof system is a pair of finite sets of formulae Γ and Δ in the language $\mathcal{L}_{T,F}$. The inductive proof system *Kripke-Feferman* (KF) consists of all the sequents, rules and axioms of the inductive system PA as in Definition 3.1, formulated in $\mathcal{L}_{T,F}$, as well as the rules and axioms below, where A is a sentence in $\mathcal{L}^{+}_{T,F}$ and r, s, t are closed terms of $\mathcal{L}^{+}_{T,F}$.

[1] This system is inspired by the finitary sequent calculus for CT given by Leigh in [Lei13], the infinitary sequent calculus for KF by Leigh in [Lei16] and the system PKF given by Halbach and Horsten in [HH06]. However, the specific representation of the sequent calculus here is our own.

Truth rules for KF[2]

$$\mathrm{L}_{T=}\ \frac{\Gamma, s^\circ = t^\circ \Rightarrow \Delta}{\Gamma, r = (s \underset{.}{=} t), Tr \Rightarrow \Delta} \qquad \mathrm{R}_{T=}\ \frac{\Gamma \Rightarrow \Delta, s^\circ = t^\circ}{\Gamma, r = (s \underset{.}{=} t) \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{T\neg=}\ \frac{\Gamma \Rightarrow \Delta, s^\circ = t^\circ}{\Gamma, r = \underset{.}{\neg}(s \underset{.}{=} t), Tr \Rightarrow \Delta} \qquad \mathrm{R}_{T\neg=}\ \frac{\Gamma, s^\circ = t^\circ \Rightarrow \Delta}{\Gamma, r = \underset{.}{\neg}(s \underset{.}{=} t) \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{T\wedge}\ \frac{\Gamma, Ts, Tt \Rightarrow \Delta}{\Gamma, r = (s \underset{.}{\wedge} t), Tr \Rightarrow \Delta} \qquad \mathrm{R}_{T\wedge}\ \frac{\Gamma \Rightarrow \Delta, Ts \qquad \Gamma \Rightarrow \Delta, Tt}{\Gamma, r = (s \underset{.}{\wedge} t) \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{T\vee}\ \frac{\Gamma, Ts \Rightarrow \Delta \qquad \Gamma, Tt \Rightarrow \Delta}{\Gamma, r = (s \underset{.}{\vee} t), Tr \Rightarrow \Delta} \qquad \mathrm{R}_{T\vee}\ \frac{\Gamma \Rightarrow \Delta, Ts, Tt}{\Gamma, r = (s \underset{.}{\vee} t) \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{T\forall}\ \frac{\Gamma, TA[t/x] \Rightarrow \Delta}{\Gamma, r = (\underset{.}{\forall} x.A), Tr \Rightarrow \Delta} \qquad \mathrm{R}_{T\forall}\ \frac{\Gamma \Rightarrow \Delta, TA[y/x]}{\Gamma, r = (\underset{.}{\forall} x.A) \Rightarrow \Delta, Tr}\ {\scriptstyle y \notin \mathrm{FV}(\Gamma, \Delta, \underset{.}{\forall} x.A)}$$

$$\mathrm{L}_{T\exists}\ \frac{\Gamma, TA[y/x] \Rightarrow \Delta}{\Gamma, r = (\underset{.}{\exists} x.A), Tr \Rightarrow \Delta}\ {\scriptstyle y \notin \mathrm{FV}(\Gamma, \Delta, \underset{.}{\exists} x.A)} \qquad \mathrm{R}_{T\exists}\ \frac{\Gamma \Rightarrow \Delta, TA[t/x]}{\Gamma, r = (\underset{.}{\exists} x.A) \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{TT}\ \frac{\Gamma, Ts \Rightarrow \Delta}{\Gamma, r = \underset{.}{T} s, Tr \Rightarrow \Delta} \qquad \mathrm{R}_{TT}\ \frac{\Gamma \Rightarrow \Delta, Ts}{\Gamma, r = \underset{.}{T} s \Rightarrow \Delta, Tr}$$

$$\mathrm{L}_{TF}\ \frac{\Gamma, Fs \Rightarrow \Delta}{\Gamma, r = \underset{.}{F} s, Tr \Rightarrow \Delta} \qquad \mathrm{R}_{TF}\ \frac{\Gamma \Rightarrow \Delta, Fs}{\Gamma, r = \underset{.}{F} s \Rightarrow \Delta, Tr}$$

Falsity rules for KF

[2]In the conclusions of the following truth and falsity rules, in addition to the principal formula of the sequent of the form Tr of Fr, we add some expression that explains the content of the term r. For example, in the rule $\mathrm{L}_T =$, we write $r = (s \underset{.}{=} t)$. This notation, which is not standard in most sequent calculus systems, is both conventional for proof systems of truth theories (see, e.g., [Lei13]), but also denotes an important distinction between a sentence $T(s \underset{.}{=} t)$ and Tr where $r = (s \underset{.}{=} t)$; deriving the former is an inference of only that specific statement, while the deriving latter implies that we may infer Tr for every term r that has the same value as $(s \underset{.}{=} t)$.

$$\mathrm{L}_{F=}\ \frac{\Gamma \Rightarrow \Delta, s^\circ = t^\circ}{\Gamma, r = (s \underset{\cdot}{=} t), Fr \Rightarrow \Delta} \qquad \mathrm{R}_{F=}\ \frac{\Gamma, s^\circ = t^\circ \Rightarrow \Delta}{\Gamma, r = (s \underset{\cdot}{=} t) \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{F\neg=}\ \frac{\Gamma, s^\circ = t^\circ \Rightarrow \Delta}{\Gamma, r = \underset{\cdot}{\neg}(s \underset{\cdot}{=} t), Fr \Rightarrow \Delta} \qquad \mathrm{R}_{F\neg=}\ \frac{\Gamma \Rightarrow \Delta, s^\circ = t^\circ}{\Gamma, r = \underset{\cdot}{\neg}(s \underset{\cdot}{=} t) \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{F\wedge}\ \frac{\Gamma, Fs \Rightarrow \Delta \qquad \Gamma, Ft \Rightarrow \Delta}{\Gamma, r = (s \underset{\cdot}{\wedge} t), Fr \Rightarrow \Delta} \qquad \mathrm{R}_{F\wedge}\ \frac{\Gamma \Rightarrow \Delta, Fs, Ft}{\Gamma, r = (s \underset{\cdot}{\wedge} t) \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{F\vee}\ \frac{\Gamma, Fs, Ft \Rightarrow \Delta}{\Gamma, r = (s \underset{\cdot}{\vee} t), Fr \Rightarrow \Delta} \qquad \mathrm{R}_{F\vee}\ \frac{\Gamma \Rightarrow \Delta, Fs \qquad \Gamma \Rightarrow \Delta, Ft}{\Gamma, r = (s \underset{\cdot}{\vee} t) \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{F\forall}\ \frac{\Gamma, FA[y/x] \Rightarrow \Delta}{\Gamma, r = (\underset{\cdot}{\forall} x.A), Fr \Rightarrow \Delta}\ {\scriptstyle y \notin \mathrm{FV}(\Gamma, \Delta, \underset{\cdot}{\forall} x.A)} \qquad \mathrm{R}_{F\forall}\ \frac{\Gamma \Rightarrow \Delta, FA[t/x]}{\Gamma, r = (\underset{\cdot}{\forall} x.A) \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{F\exists}\ \frac{\Gamma, FA[y/x] \Rightarrow \Delta}{\Gamma, r = (\underset{\cdot}{\exists} x.A), Fr \Rightarrow \Delta} \qquad \mathrm{R}_{F\exists}\ \frac{\Gamma \Rightarrow \Delta, FA[y/x]}{\Gamma, r = (\underset{\cdot}{\exists} x.A) \Rightarrow \Delta, Fr}\ {\scriptstyle y \notin \mathrm{FV}(\Gamma, \Delta, \underset{\cdot}{\exists} x.A)}$$

$$\mathrm{L}_{FT}\ \frac{\Gamma, Fs \Rightarrow \Delta}{\Gamma, r = \underset{\cdot}{T} s, Fr \Rightarrow \Delta} \qquad \mathrm{R}_{FT}\ \frac{\Gamma \Rightarrow \Delta, Fs}{\Gamma, r = \underset{\cdot}{T} s \Rightarrow \Delta, Fr}$$

$$\mathrm{L}_{FF}\ \frac{\Gamma, Ts \Rightarrow \Delta}{\Gamma, r = \underset{\cdot}{F} s, Fr \Rightarrow \Delta} \qquad \mathrm{R}_{FF}\ \frac{\Gamma \Rightarrow \Delta, Ts}{\Gamma, r = \underset{\cdot}{F} s \Rightarrow \Delta, Fr}$$

Axioms for KF

$$\Gamma, Tr \Rightarrow \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(r) \quad \Gamma, Fr \Rightarrow \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(r)$$

The rules of the above system translate KF's axioms from Definition 2.13 into sequent calculus rules. They furthermore derive all the axioms of the theory. Therefore, any soundness and completeness results that hold for the standard theory KF extend to equivalent results in the inductive proof system.

4.2 Minimality and maximality in the KF proof system

4.2.1 The system μKF

The theory μKF as devised by Burgess in [Bur14] contains the axioms KF1-KF8 in the compositional direction, as well as KF9, but also includes a minimality scheme. This allows

μKF to be presented as an axiomatization of Kripke's least fixed point model. Since the purpose of this thesis is to develop a cyclic system for KF, which does not specify any property of the truth predicate as a fixed point, it will be interesting to also pay attention to μKF and how cyclic argumentation over the least fixed point looks as a specific case.

Definition 4.2 (μKF)**.** The proof system μKF consists of all the sequents, rules and axioms of KF as in Definition 4.1 and is extended with the minimality axiom:

$$\Rightarrow M(P, Q) \to \forall x.((Tx \to P(x)) \wedge (Fx \to Q(x)))$$

for all formulae $P(x), Q(x)$ of $\mathcal{L}^+_{T,F}$, where M is as defined in Definition 2.15.

As in the case of the proof system KF, since μKF's inductive proof system characterizes all axioms of the theory μKF, any soundness and completeness results the hold for the theory extend to equivalent results for the proof system.

4.2.2 The greatest fixed point model

Kripke's original outline of a theory did not consider the greatest fixed point model in his construction for truth, and, as far as the author is aware, no further attempt at developing such a model, nor an axiomatic theory for it, has been made since. However, for our purposes of developing a cyclic proof theory of KF, it might be interesting to construct the greatest fixed point model and a proof-theoretic representation that goes along with it, to see how it behaves differently from μKF and KF. We construct the greatest fixed point model, which we call $\mathfrak{G}$, in a similar but dual way to Kripke's least fixed point model $\mathfrak{K}$. We will briefly outline of the construction of $\mathfrak{G}$ before going on to give the proof system νKF.

We define a model $\mathfrak{G}^\alpha$ in the greatest fixed point construction as

$$\mathfrak{G}^\alpha := (\mathbb{N}, \Lambda^\alpha(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}), \bar{\Lambda}^\alpha(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}))$$

where Λ and $\bar{\Lambda}$ are as defined for Definition 2.7 in Chapter 2, applied to the set of all sentences in the language $\mathcal{L}^+_{T,F}$. Specifically, $\Lambda^\alpha(\mathrm{Sent}_{\mathcal{L}^+_{T,F}})$ denotes the α-th iteration of the operator Λ, such that:

$$\Lambda^0(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}) := \mathrm{Sent}_{\mathcal{L}^+_{T,F}} \qquad \Lambda^{\alpha+1}(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}) := \Lambda(\Lambda^\alpha(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}))$$

$$\Lambda^\lambda(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}) := \bigcap_{\alpha<\lambda} \Lambda^\alpha(\mathrm{Sent}_{\mathcal{L}^+_{T,F}})$$

for λ a limit ordinal. $\bar{\Lambda}^\alpha$ is defined similarly.

As in the positive inductive construction for the Kripke least fixed point model, by the Knaster-Tarski theorem we know that there is an ordinal θ such that for every $\alpha \geqslant \theta$ $\mathfrak{G}^{\alpha+1} = \mathfrak{G}^\alpha$. Thus the *fixed point models* of this construction are all models $\mathfrak{G}^\alpha$ with $\alpha \geqslant \theta$ for θ being a closure ordinal. We then take the *greatest fixed point model* in this construction to be $\mathfrak{G} := \mathfrak{G}^\theta = (\mathbb{N}, \Lambda^\theta(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}), \bar{\Lambda}^\theta(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}))$.

4.2.3 The system νKF

We now construct the theory νKF, which is obtained from KF by extending it with a scheme for maximality; this theory is essentially "dual" to μKF. Since νKF occur in previous literature, in this thesis it is mostly used as a tool in the same way as μKF; it functions as a comparison to KF, and specifically when developing cyclic systems for these theories we shall see how thir differing properties affect the progressivity conditions of these systems.

The theory νKF consists of the decompositional (meaning the left-to-right) direction of axioms KF1-KF8, as well as KF9, with an added scheme for maximality. If we let $N(P, Q)$ be the dual counterpart to the minimality sentence $M(P, Q)$ described in Definition 2.15, then N is constructed as follows:

Definition 4.3. We define N as: $N(P(x), Q(x)) :=$

$$\forall x.\forall y.((P(x \underset{\cdot}{=} y) \to x^\circ = y^\circ))\wedge$$

$$\forall x.\forall y.((Q(x \underset{\cdot}{=} y) \to \neg x^\circ = y^\circ))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\wedge} y) \to (P(x \underset{\cdot}{\wedge} y) \to P(x) \wedge P(y)))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\wedge} y) \to (Q(x \underset{\cdot}{\wedge} y) \to Q(x) \vee Q(y)))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\vee} y) \to (P(x \underset{\cdot}{\vee} y) \to P(x) \vee P(y)))\wedge$$

$$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\vee} y) \to (Q(x \underset{\cdot}{\vee} y) \to Q(x) \wedge Q(y)))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} A.x) \to (P(\underset{\cdot}{\forall} A.x) \to \forall y.P(A[y/x])))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} A.x) \to (Q(\underset{\cdot}{\forall} A.x) \to \exists y.Q(A[y/x])))\wedge$$

$$\forall A.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\exists} A.x) \to (P(\underset{\cdot}{\exists} A.x) \to \exists y.P(A[y/z])))\wedge$$

$$\forall y.\forall x.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\exists} A.x) \to (Q(\underset{\cdot}{\exists} A.x) \to \forall y.Q(A[y/x])))\wedge$$

$$\forall t.(P(\underset{\cdot}{T}t) \to P(t^\circ)) \wedge \forall t.(P(\underset{\cdot}{F}t) \to Q(t^\circ)))\wedge$$

$$\forall t.(Q(\underset{\cdot}{T}t) \to Q(t^\circ)) \wedge \forall t.(Q(\underset{\cdot}{F}t) \to P(t^\circ)))\wedge$$

$$\forall x.((P(x) \wedge Q(x)) \to \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x))$$

The proof system νKF is then defined as:

Definition 4.4 (νKF)**.** The proof system νKF consists of all the sequents, rules and axioms of KF as in Defintion 4.1, and is extended with the maximality axiom:

$$\textsc{Max} \;\; \Rightarrow N(P, Q) \rightarrow \forall x.((P(x) \rightarrow Tx) \wedge (Q(x) \rightarrow Fx))$$

for all formulae $P(x), Q(x)$ of $\mathcal{L}^{+}_{T,F}$, where N is as defined in Definition 4.3.

Given inductive proof systems for these three theories, we are now ready to explore their cyclic proof-theoretic counterparts.

Chapter 5

Cyclic Proof-Theoretic Representations of KF

In this chapter, we propose a cyclic proof system for KF, called CKF. In addition to this we introduce auxiliary cyclic proof systems for the theories μKF and νKF, which model truth as a least and as a greatest fixed point respectively. These systems, called CμKF and CνKF, therefore extend CKF by specifying which fixed point T and F are identified as, which we then cycle over.

For all three systems, we formulate appropriate progressivity conditions on the truth and falsity predicates that allow us to both get a sound proof, and embed their inductive counterparts. We will discuss some implications of these results in Chapter 6.

5.1 Rules of cyclic systems for KF, μKF and νKF

The cyclic system CKF consists of the same rules as the finitary sequent calculus of KF, but with the induction rule replaced by $\textsc{Case}_x$ and $\textsc{Subst}$, as defined in Chapter 3 for CPA.

Definition 5.1 (Cyclic Kripke-Feferman)**.** The proof system *cyclic Kripke-Feferman* (CKF) consists of all the sequents, rules and axioms of KF as in Definition 4.1, except for the induction axiom, but extended with the following rules:

$$\textsc{Subst}\ \frac{\Gamma \Rightarrow \Delta}{\Gamma[t/x] \Rightarrow \Delta[t/x]}\ x \notin \mathrm{FV}(t)$$

$$\textsc{Case}_x\ \frac{\Gamma[0/x] \Rightarrow \Delta[0/x] \qquad \Gamma[s(x)/x] \Rightarrow \Delta[s(x)/x]}{\Gamma \Rightarrow \Delta}\ x \text{ is fresh}$$

The cyclic systems CμKF and CνKF will contain the same axioms and rules of KF, and will be enriched with a progressivity condition for the truth and falsity rules that expresses minimality of truth as a fixed point. We will show that the minimality axiom as given in Definition 4.2, which is retained in the inductive proof system, will be subsumed by cyclic reasoning following from this progressivity condition within CμKF.

Definition 5.2 (Cyclic μKF)**.** The proof system *cyclic* μKF (CμKF) consists of all the sequents, rules and axioms of CKF as in Definition 5.1.

Definition 5.3 (Cyclic νKF)**.** The proof system *cyclic* νKF (CνKF) consists of all the sequents, rules and axioms of CKF as in Definition 5.1.

5.2 Proof in cyclic KF, μKF and νKF

We know that CμKF and CνKF contain all the same inference rules and axioms as CKF. Therefore, pre-proofs are defined the same in all three systems:

Definition 5.4 (Pre-proof)**.** A *pre-proof* in CKF, CμKF or CνKF is a pair $\Pi = (C, \lambda)$ where C is a cyclic tree, and λ is a labelling function $\lambda : Node(P) \to Seq$ labelling the nodes of P with sequents, such that each inner node is labelled by one of the CKF derivation rules, leaves are labelled by axioms of CKF and buds share a label with their companion.

We also define the notion of a *truth trace* as a trace along the truth and falsity rules in a derivation, which is the same in all three systems:

Definition 5.5 (Truth trace)**.** A sequence of formulae $(A_i)_{i\in\omega}$ is a *truth trace* along an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ if

- There exists an $n \in \omega$ such that A_i is a formula in $\Gamma_i \Rightarrow \Delta_i$ for all $i \geqslant n$;
- $A_{i+1} = A_i$ if A_i is not the principal formula in $\Gamma_i \Rightarrow \Delta_i$, otherwise A_{i+1} is the residual subformula of A_i in the premise of the sequent $\Gamma_i \Rightarrow \Delta_i$.

This definition will mainly play a part when defining proof in CμKF and CνKF. The global progressivity condition is where the main difference between the systems CμKF, CνKF or CKF lies, due to the fact that each system takes truth as a different type of fixed point (minimal, maximal, or arbitrary). In the next subsections we will define when a trace is progressing for the system CKF, which we then extend to define progressing traces and proofs in CμKF and its dual CνKF

5.2.1 Proof and progressivity in CKF

In contrast to the systems CμKF and CνKF, we will design the system CKF to model the truth and falsity predicates as any arbitrary fixed point, i.e., we do not restrict it to be either the least or the greatest. Since cyclic proof systems are typically defined in relation to specified least or greatest fixed points, it is less straightforward to identify the appropriate progressivity condition for a theory of arbitrary fixed points. However, CKF already contains the theory CPA. Therefore, we can extend the underlying progressivity condition of CPA to also apply to the rest of CKF. In fact, this suffices as a progressivity condition to obtain sound proofs in CKF. Since there are no criteria for specifying fixed points in CKF, any progressivity condition on truth traces can be disregarded. We can therefore define a proof in CKF with the following progressivity condition:

Definition 5.6 (Progressing trace in CKF)**.** A trace in CKF is *progressing* if it is an arithmetic trace as in Definition 3.5 and is infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x.

Definition 5.7 (CKF proof)**.** A pre-proof in CKF Π is a *proof* in CKF if along every infinite branch $\Gamma_i \Rightarrow \Delta_i$ there exists some progressing trace along the branch.

Embedding KF into CKF

When we adhere to this definition of a proof in CKF we can derive the induction axiom of KF for any formula $A(x)$ in $\mathcal{L}_{T,F}$. For example, the following proof (which is essentially identical to the proof tree in Figure 3.3) is a proof in CKF, where $A(x) := Tx$ and Γ abbreviates $T(0), \forall x.(Tx \to T(s(x))$:

$$
\textsc{Subst}\,\dfrac{\textsc{Case}_x\,\dfrac{\mathrm{Ax}\,\dfrac{}{\Gamma \Rightarrow T(0)} \qquad \mathrm{L}_\forall\,\dfrac{\mathrm{L}_\to\,\dfrac{\mathrm{Ax}\,\dfrac{}{\Gamma, T(s(x)) \Rightarrow T(s(x))} \qquad \mathrm{Wk}\,\dfrac{\Gamma \Rightarrow Tx(\star)}{\Gamma \Rightarrow T(s(x)), Tx}}{\Gamma, Tx \to T(s(x)) \Rightarrow T(s(x))}}{\Gamma \Rightarrow T(s(x))}}{\Gamma \Rightarrow Tx(\star)}}{\Gamma \Rightarrow Tt}
$$

Figure 5.1: Derivation of the induction axiom for T in CKF

By how we have defined CKF, the only aspect where the rules of CKF differ from those of inductive KF is in the induction axiom, which is replaced by the cyclic rules $\textsc{Subst}$ and $\textsc{Case}_x$ in CKF. In Theorem 3.10, we have shown that these rules can simulate the induction axiom. Consequently, as demonstrated in and Figure 5.1, the embedding of PA into CPA given in Theorem 3.10 smoothly lifts to an embedding of KF into CKF, which means that induction can be subsumed by cyclic reasoning in this system. Therefore, we have:

Theorem 5.8. KF $\subseteq$ CKF

Thus any theorem provable in KF is provable in CKF. Furthermore, we will now demonstrate how proofs in CKF that adhere to this condition are sound with respect to the Kripke fixed point model for an arbitrary fixed point, and that this also suffices to show that the inductive system KF can be embedded in CKF.

Soundness of CKF

For the sake of giving a soundness argument for CKF, we will employ the positive inductive definition of $\mathfrak{K}$ as in Definition 2.7. By Corollary 2.9, any fixed point model, denoted here as $\mathfrak{K}'$, arising from this approach is equivalent to a fixed point model in Kripke's original approach, so no issues arises in choosing this construction over the other, but the positive inductive definition will allow us to derive the lemma below, which is relevant for obtaining soundness results. Since the theories μKF, νKF and KF are formulated in classical logic, our soundness argument for each of these theories will be based on classical reasoning over the relevant structure. To obtain soundness results for any of our cyclic systems we must first derive the following lemma.

Lemma 5.9 (Reflecting falsity)**.** *Let $\mathfrak{K}'$ be any fixed point model of the Kripke construction. Given an inference step, where* r *is one of the inference rules for* CKF*:*

$$\mathsf{r}\,\frac{\Gamma_0 \Rightarrow \Delta_0 \qquad \cdots \qquad \Gamma_n \Rightarrow \Delta_n}{\Gamma \Rightarrow \Delta}$$

If $\mathfrak{K}', \rho \models \bigwedge \Gamma$ and $\mathfrak{K}', \rho \not\models \bigvee \Delta$ for some environment $\rho : FV(\Gamma) \cup FV(\Delta) \to \omega$ interpreting the free variables of any formula $A \in \Gamma \cup \Delta$ as natural numbers then there is $\rho' \leqslant \rho$ and $0 \leqslant i \leqslant n$ such that $\mathfrak{K}', \rho' \models \bigwedge \Gamma_i$ and $\mathfrak{K}', \rho' \not\models \bigvee \Delta_i$.

Proof. In the case that r is any of the arithmetic rules coming from CPA, we are not interested in $\mathfrak{K}'$ and only look at the environment ρ. Therefore in these cases the proof follows the same structure as the proof of Lemma 3.7 and if $\mathfrak{K}', \rho \not\models \Gamma \Rightarrow \Delta$ then we have $\mathfrak{K}', \rho' \not\models \Gamma_i \Rightarrow \Delta_i$, where $\Gamma_i \Rightarrow \Delta_i$ is some premise of the conclusion.

In the case r is one of the truth and falsity rules, the environment ρ remains unchanged, so we have $\rho' = \rho$ and if the conclusion of the rule does not hold in ρ, neither will the premise. Furthermore, by the positive inductive definition of $\mathfrak{K}'$ and the fact it identifies a fixed point model, so therefore Corollary 2.6 holds, for any truth or falsity formula $A \in \Gamma \cup \Delta$, iff A is in $\mathfrak{K}'$ then its residual subformula A_i is in $\mathfrak{K}'$. Thus, in the case of a truth or falsity rule, if $\mathfrak{K}', \rho \not\models \Gamma \Rightarrow \Delta$ then we have $\mathfrak{K}', \rho \not\models \Gamma_i \Rightarrow \Delta_i$, where $\Gamma_i \Rightarrow \Delta_i$ is some premise of the conclusion. □

From this lemma we also get the following corollary:

Corollary 5.10. *Let* d *be a be a pre-proof with endsequent $\Gamma \Rightarrow \Delta$ and let $\mathfrak{K}'$ be any fixed point model of the Kripke construction. If $\mathfrak{K}', \rho \not\models \Gamma \Rightarrow \Delta$ then we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i \in \omega}$ through* d *and a family of environments $(\rho_i)_{i \in \omega}$, with $\rho_{i+1} \leqslant \rho_i$ such that $\mathfrak{K}', \rho_i \not\models \Gamma_i \Rightarrow \Delta_i$ for all $i \in \omega$.*

We now show a soundness result for the system CKF, demonstrating the system is sound with respect to any fixed point model of the positive inductive Kripke construction. In fact, this follows almost directly from the soundness of CPA, but must be slightly adapted to the Kripke fixed point model, which is where Lemma 5.9 and Corollary 5.10 comes in.

Theorem 5.11 (Soundness)**.** *If* CKF $\vdash \Gamma \Rightarrow \Delta$ *then* $\mathfrak{K}' \models \Gamma \Rightarrow \Delta$ *for* $\mathfrak{K}'$ *is any fixed point model for the Kripke construction.*

Proof. Let Π be a CKF proof of $\Gamma \Rightarrow \Delta$ and suppose towards contradiction that $\mathfrak{K}' \not\models \Gamma \Rightarrow \Delta$. By Corollary 5.10, we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ and a corresponding family of evironments $(\rho_i)_{i\in\omega}$ falsifying each sequent $\Gamma_i \Rightarrow \Delta_i$. By the progressivity condition of Π we know that this infinite branch has a progressing trace. This means that, by Definition 5.6, the trace is arithmetic and infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x. As observed in Lemma 3.7, whenever $\rho_i \not\models \Gamma \Rightarrow \Delta$, and we pass the right-hand side of $\textsc{Case}_x$, then $\rho_{i+1} \not\models \Gamma \Rightarrow \Delta$ with $\rho_{i+1}(x) < \rho_i(x)$. Since we pass this rule infinitely often, we have $\rho_{i+1}(x) < \rho_i(x)$ for infinitely many i, contradicting the well-foundedness of the natural numbers. Thus $\mathfrak{K}' \models \Gamma \Rightarrow \Delta$ has to hold for any provable sequent $\Gamma \Rightarrow \Delta$ in Π. $\square$

5.2.2 Proof and progressivity in CμKF

It has already been proven by Burgess [Bur14, p. 19] that KF is a subtheory of μKF, so it would make sense for the cyclic version of μKF to contain the cyclic version of KF. Thus, CμKF should have a progressing trace definition that allows it to be an extension of CKF. In addition to this, and to giving a sound proof system, the goal of the progressing truth traces in CμKF and CνKF is to allow us to embed the inductive minimality or maximality axioms.

For these reasons, in CμKF and CνKF, we propose that the cyclic argumentation can come from either the arithmetic progressivity condition, as laid out for CKF, or from some additional progressivity condition that cycles over the truth and falsity rules. We define for CμKF a progressing trace such that it allows us to both derive sound proofs, and to directly simulate the inductive minimality scheme of μKF into CμKF.

Since μKF models truth as a least fixed point, and therefore as an inductively defined object, by the positive inductive construction of the Kripke least fixed point model $\mathfrak{K}$, every proof of a truth formula TA (or a falsity formula FA) ought to provide evidence that A belongs to S (resp., $\bar{S}$), i.e. the truth set (resp., the falsity set) of $\mathfrak{K}$. Consequently, such a proof must be *finite*; it can be derived in our proof system only through finite applications of right-hand side rules for T and F. Therefore, any infinitary cyclic argumentation can only occur on the left-hand side of the sequent. Keeping this intuition in mind, we define our progressivity condition as follows:

Definition 5.12 (Progressing trace in CμKF)**.** A trace in CμKF is *progressing* if it is either an arithmetic trace as in Definition 3.5 and is infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x; or it is a truth trace as in Definition 5.5 and infinitely often principal for any rule of the form L_{T*} or L_{F*}, where $* \in \{\wedge, \vee, \forall, \exists, T, F\}$.

Definition 5.13 (CμKF proof)**.** A pre-proof in CμKF Π is a *proof* in CμKF if along every infinite branch $\Gamma_i \Rightarrow \Delta_i$ there exists some progressing trace along the branch.

Thus a progressing trace will have to pass through either $\textsc{Case}_x$, or a left truth or falsity rule (with the exception of the left truth and falsity rules for (in)equality) infinitely often.

The following is a proof in CμKF of some truth statement on the left:

$$\textsc{Cut}\,\dfrac{\textsc{Wk}\,\dfrac{\textsc{Ax}\,\dfrac{}{A \Rightarrow A}\,(A \in \Gamma)}{\Gamma, Tz \Rightarrow \Delta, A, z = \underset{\cdot}{T}s} \qquad \mathrm{L}_{TT}\,\dfrac{\textsc{Subst}\,\dfrac{\Gamma, Tz \Rightarrow \Delta, A(\star)}{\Gamma, Ts \Rightarrow \Delta, A}}{\Gamma, z = \underset{\cdot}{T}s, Tz \Rightarrow \Delta, A}}{\Gamma, Tz \Rightarrow \Delta, A(\star)}$$

Figure 5.2: Example proof in CμKF

Note that while this is a valid proof in CμKF, it is not a proof in CKF, since it does not satisfy CKF's requirement of having a progressing arithmetic trace; there is only a truth trace which is progressing according to CμKF's definition.

Embedding μKF into CμKF

We show that adding the above progressing trace condition to our proof system CμKF allows us to get rid of the inductive minimality axiom, since any proof in μKF with this scheme can be simulated by a proof in CμKF. Thus for any proof of a sequent in μKF. there is a proof of the same sequent in CμKF.

Theorem 5.14. $\mu\mathrm{KF} \subseteq \mathrm{C}\mu\mathrm{KF}$

Proof. We have to show that in a derivation of $\textsc{Min}$ as in Definition 4.2 any infinite trace is progressing and therefore the derivation is a proof in CμKF. The prooftree of this derivation begins at the root as follows:

$$\mathrm{R}_{\to}\,\dfrac{\mathrm{R}_{\forall}\,\dfrac{\mathrm{R}_{\wedge}\,\dfrac{\mathrm{R}_{\to}\,\dfrac{\begin{array}{c}\vdots\\ M, Tz \Rightarrow P(z)(\star)\end{array}}{M \Rightarrow (Tz \to P(z))} \qquad \mathrm{R}_{\to}\,\dfrac{\begin{array}{c}\vdots\\ M, Fz \Rightarrow Q(z)(\blacktriangle)\end{array}}{M \Rightarrow Fz \to Q(z)}}{M \Rightarrow Tz \to P(z)) \wedge (Fz \to Q(z)}}{M \Rightarrow \forall x.((Tx \to P(x)) \wedge (Fx \to Q(x))}}{\Rightarrow M \to \forall x.((Tx \to P(x)) \wedge (Fx \to Q(x)))}$$

Figure 5.3: The beginning of the derivation.

The inner nodes marked with $(\star)$ and $(\blacktriangle)$ denote that these nodes are companion nodes for leaves in the continuation of the proof marked with the same symbol. Given that z

must be a sentence since it occurs in Tz, Fz and $P(z), Q(z)$, it can take several forms. We can therefore take $\bigvee z$ to be the disjunction:

$$\bigvee z := \exists x.\exists y.(z = (x \dot{=} y)) \vee \exists x.\exists y.(z = (x \dot{\wedge} y)) \vee \exists x.\exists y.(z = (x \dot{\vee} y)) \vee \exists x \exists$$
$$A(z = (\dot{\forall} x.A)) \vee \exists x.\exists A.(z = (\dot{\exists} x.A)) \vee \exists y.(z = (\dot{T} y)) \vee \exists y.(z = ((\dot{F})y))$$

where x, y denote the codes of a term. From the definition of a sentence, we know it is possible to construct a proof in PA of the sequent

$$\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z) \Rightarrow \bigvee z$$

By Theorem 3.9, stating the equivalence of PA and CPA, this inductive argument can be substituted by a cyclic proof in CPA, thus the sequent is derivable in our cyclic system for CμKF as well since it contains CPA.

We use this information to apply a Cut-rule, and then we apply the $\mathrm{L}_\vee$-rule (multiple times) to decompose the disjunction $\bigvee z$:

$$\dfrac{\begin{array}{c}\vdots\\ \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z) \Rightarrow \bigvee z\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\\ z = (s \dot{=} t), M, Tz \Rightarrow P(z)\end{array}}{\exists x.\exists y.(z = (x \dot{=} y)), M, Tz \Rightarrow P(z)}\,\mathrm{L}_\exists,\mathrm{L}_\exists \qquad \dfrac{\begin{array}{c}\vdots\\ z = (s \dot{\wedge} t), M, Tz \Rightarrow P(z)\end{array}}{\exists x.\exists y.(z = (x \dot{\wedge} y)), M, Tz \Rightarrow P(z)}\,\mathrm{L}_\exists,\mathrm{L}_\exists \qquad \begin{array}{c}\vdots\\ \text{other disjuncts of } z\\ \vdots\\ \vdots\end{array}}{\bigvee z, M, Tz \Rightarrow P(z)}\,\mathrm{L}_\vee,\mathrm{L}_\vee,\ldots}{M, Tz \Rightarrow P(z)(\star)}\,\textsc{Cut}$$

Figure 5.4: Applying Cut, $\mathrm{L}_\vee$ and $\mathrm{L}_\exists$

Likewise if we instead choose to follow the right-hand branch. Next, on each branch, we apply an $\mathrm{L}_\wedge$ rule to decompose M, choosing the suitable conjunct that matches the operator of z as our principal formula along which to develop a trace. We look at some of the branches and omit the remaining cases. Let's for example follow the branch of the above derivation that takes the disjunct $z = (s \dot{\wedge} t)$ and follow its trace. For brevity we omit some of the non-principal formulae of the proof, for instance the non-principal subformulae of M. We continue the tree in Figure 5.4 along the left branch as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{M, T(z) \Rightarrow P(z), \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z)}\,\mathrm{Ax}}{M, z = (s \dot{\wedge} t), T(z) \Rightarrow P(z), \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(s \dot{\wedge} t)}\,\mathrm{L}_= \qquad \dfrac{\dfrac{\dfrac{}{P(z), M, T(z) \Rightarrow P(z)}\,\mathrm{Ax}}{P(s \dot{\wedge} t), M, z = (s \dot{\wedge} t), T(z) \Rightarrow P(z)}\,\mathrm{L}_= \quad \text{(cont'd in Figure 5.6)}}{(P(s) \wedge P(t) \to P(s \dot{\wedge} t)), z = (s \dot{\wedge} t), M, T(z) \Rightarrow P(z)}\,\mathrm{L}_\to}{\dfrac{\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(s \dot{\wedge} t) \to (P(s) \wedge P(t) \to P(s \dot{\wedge} t)), z = (s \dot{\wedge} t), M, T(z) \Rightarrow P(z)}{\dfrac{\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \dot{\wedge} y) \to (P(x) \wedge P(y) \to P(x \dot{\wedge} y)), z = (s \dot{\wedge} t), M, T(z) \Rightarrow P(z)}{z = (s \dot{\wedge} t), M, Tz \Rightarrow P(z)}\,\mathrm{L}_\wedge}\,\mathrm{L}_\forall,\mathrm{L}_\forall}\,\mathrm{L}_\to$$

Figure 5.5: $z = (s \dot{\wedge} t)$ (1)

$$
\begin{prooftree}
\AxiomC{$M, Tz \Rightarrow P(z)(\star)$}
\LeftLabel{\textsc{Subst}}
\UnaryInfC{$M, Ts \Rightarrow P(s)$}
\LeftLabel{\textsc{Wk}}
\UnaryInfC{$M, z = (s \underset{\cdot}{\wedge} t), Ts, Tt \Rightarrow P(z), P(s)$}
\LeftLabel{$\mathrm{L}_{T\wedge}$}
\UnaryInfC{$z = (s \underset{\cdot}{\wedge} t), M, T(z) \Rightarrow P(z), P(s)$}
\AxiomC{$M, Tz \Rightarrow P(z)(\star)$}
\LeftLabel{\textsc{Subst}}
\UnaryInfC{$M, Tt \Rightarrow P(t)$}
\LeftLabel{\textsc{Wk}}
\UnaryInfC{$M, z = (s \underset{\cdot}{\wedge} t), Ts, Tt \Rightarrow P(z), P(t)$}
\LeftLabel{$\mathrm{L}_{T\wedge}$}
\UnaryInfC{$z = (s \underset{\cdot}{\wedge} t), M, T(z) \Rightarrow P(z), P(t)$}
\LeftLabel{$\mathrm{R}_\wedge$}
\BinaryInfC{$z = (s \underset{\cdot}{\wedge} t), M, T(z) \Rightarrow P(z), P(s) \wedge P(t)$}
\end{prooftree}
$$

Figure 5.6: $z = (s \underset{\cdot}{\wedge} t)$ (2)

The nodes marked with $(\star)$ denote buds, whose companion is marked with the same symbol in Figure 5.3. The prooftree expands similarly when taking any other disjunct of $\bigvee z$, and choosing the suitable conjunct of M as principal formula, as well as when we are following the right branch of the initial tree, which unfolds the F-predicate. Let's also look at the case $z = \forall x.A$:

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$Tz \Rightarrow P(z), \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z)$}
\LeftLabel{$\mathrm{L}_=$}
\UnaryInfC{$z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z), \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} x.A)$}
\AxiomC{(cont'd in Figure 5.8)}
\LeftLabel{$\mathrm{L}_\to$}
\BinaryInfC{$(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} x.A) \to (\forall y.P(A[y/x]) \to P(\underset{\cdot}{\forall} x.A)), z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z)$}
\LeftLabel{$\mathrm{L}_\forall$, $\mathrm{L}_\forall$}
\UnaryInfC{$\forall x.\forall A.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(\underset{\cdot}{\forall} x.A) \to (\forall y.P(A[y/x]) \to P(\underset{\cdot}{\forall} x.A)), z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z)$}
\LeftLabel{$\mathrm{L}_\wedge$}
\UnaryInfC{$z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z)$}
\end{prooftree}
$$

Figure 5.7: $z = (\underset{\cdot}{\forall} x.A)$ (1)

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$P(z), M, Tz \Rightarrow P(z)$}
\LeftLabel{$\mathrm{L}_=$}
\UnaryInfC{$P(\underset{\cdot}{\forall} x.A)), z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z)$}
\AxiomC{$M, Tz \Rightarrow P(z)(\star)$}
\LeftLabel{\textsc{Subst}}
\UnaryInfC{$M, TA[t/x], \Rightarrow, P(A[t/x])$}
\LeftLabel{\textsc{Wk}}
\UnaryInfC{$z = (\underset{\cdot}{\forall} x.A), M, TA[t/x], \Rightarrow P(z), P(A[t/x])$}
\LeftLabel{$\mathrm{L}_{T\forall}$}
\UnaryInfC{$z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z), P(A[t/x])$}
\LeftLabel{$\mathrm{R}_\forall$}
\UnaryInfC{$z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z), \forall y.P(A[y/x])$}
\LeftLabel{$\mathrm{L}_\to$}
\BinaryInfC{$(\forall y.P(A[y/x]) \to P(\underset{\cdot}{\forall} x.A)), z = (\underset{\cdot}{\forall} x.A), M, Tz \Rightarrow P(z)$}
\end{prooftree}
$$

Figure 5.8: $z = (\underset{\cdot}{\forall} x.A)$ (2)

The only situation where we could move between the left-hand T-branches and the right-hand F-branches of Figure 5.3 is in the application of the L_{TF}- or L_{FT}-rule. Let's take the case $z = \underset{\cdot}{F} s$ for example:

$$
\mathrm{L}_{\wedge}\dfrac{\mathrm{L}_{\forall}\dfrac{\mathrm{L}_{\to}\dfrac{\mathrm{L}_{TF}\dfrac{\mathrm{W{\small K}}\dfrac{\mathrm{S{\small UBST}}\dfrac{M, Fz \Rightarrow Q(z)(\blacktriangle)}{M, Fs \Rightarrow Q(s^\circ)}}{z = \underset{\cdot}{F}s, M, Fs \Rightarrow P(z), Q(s^\circ)}}{z = \underset{\cdot}{F}s, M, Tz \Rightarrow P(z), Q(s^\circ)} \qquad \mathrm{L}_{=}\dfrac{\mathrm{A{\small X}}\dfrac{}{P(z), M, Tz \Rightarrow P(z)}}{P(\underset{\cdot}{F}s)), z = \underset{\cdot}{F}s, M, Tz \Rightarrow P(z)}}{Q(s^\circ) \to P(\underset{\cdot}{F}s), z = \underset{\cdot}{F}s, M, Tz \Rightarrow P(z)}}{\forall t.(Q(t^\circ) \to P(\underset{\cdot}{F}t)), z = \underset{\cdot}{F}s, M, Tz \Rightarrow P(z)}}{z = \underset{\cdot}{F}s, M, Tz \Rightarrow P(z)}
$$

Figure 5.9: $z = \underset{\cdot}{F}s$

Now the node marked ($\blacktriangle$) cycles back to the node marked with the same symbol of the initial tree in Figure 5.3.

In the case that z describes an (in)equality between terms, it passes the rules $\mathrm{L}_{T=}$, $\mathrm{L}_{T\neg=}$, $\mathrm{L}_{F=}$, $\mathrm{L}_{F\neg=}$ and the branch will be finite. Look for example at the case where $z = (s \underset{\cdot}{=} t)$:

$$
\mathrm{L}_{\wedge}\dfrac{\mathrm{L}_{\forall}\dfrac{\mathrm{L}_{\to}\dfrac{\mathrm{L}_{=}\dfrac{\mathrm{A{\small X}}\dfrac{}{P(z), M, Tz \Rightarrow P(z)}}{P(s \underset{\cdot}{=} t), z = (s \underset{\cdot}{=} t), M, Tz \Rightarrow P(z)} \qquad \mathrm{L}_{T=}\dfrac{\mathrm{A{\small X}}\dfrac{}{z = (s \underset{\cdot}{=} t), M, s^\circ = t^\circ \Rightarrow P(z), s^\circ = t^\circ}}{z = (s \underset{\cdot}{=} t), M, Tz \Rightarrow P(z), s^\circ = t^\circ}}{s^\circ = t^\circ \to P(s \underset{\cdot}{=} t), z = (s \underset{\cdot}{=} t), M, Tz \Rightarrow P(z)}}{\forall x.\forall y.(x^\circ = y^\circ \to P(x \underset{\cdot}{=} y)), z = (s \underset{\cdot}{=} t), M, Tz \Rightarrow P(z)}}{z = (s \underset{\cdot}{=} t), M, Tz \Rightarrow P(z)}
$$

Figure 5.10: $z = (s \underset{\cdot}{=} t)$

The remaining branches unfold similarly to the ones above. This shows that the minimality scheme can be simulated by a derivation in CμKF. Notice that the derivation is progressing: in any infinite branch we can find a trace that passes through some L_{T*} or L_{F*} rule infinitely often. Thus the derivation is a CμKF proof.

$\square$

From this proof it is apparent that the embedding of μKF into CμKF requires progressing traces of CμKF to be principal infinitely often for *any* left truth or falsity rules (with the exception of $\mathrm{L}_{T=}$, $\mathrm{L}_{T\neg=}$ and $\mathrm{L}_{F=}$, $\mathrm{L}_{F\neg=}$). Any less permissive version of the progressivity condition, for example the progressivity condition of CKF would make it impossibly to derive the minimality scheme in CμKF.

Soundness of CμKF

Using Lemma 5.9 and Corollary 5.10, we can show that any proof in CμKF is sound with respect to the least fixed point model $\mathfrak{K}$, so any sentence that is provable in CμKF is satisfiable in $\mathfrak{K}$ (according to classical logic). This proof expands on the soundness proof for CKF, taking into account the additional progressivity condition for truth traces.

Theorem 5.15 (Soundness)**.** *If* $\mathrm{C}\mu\mathrm{KF} \vdash \Gamma \Rightarrow \Delta$ *then* $\mathfrak{K} \models \Gamma \Rightarrow \Delta$

Proof. Let Π be a CμKF proof of $\Gamma \Rightarrow \Delta$ and suppose towards contradiction that $\mathfrak{K} \not\models \Gamma \Rightarrow \Delta$. By Corollary 5.10, we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ and a corresponding family of evironments $(\rho_i)_{i\in\omega}$ falsifying each sequent $\Gamma_i \Rightarrow \Delta_i$. By the progressivity condition of Π we know that the infinite branch has a progressing trace. This means that, by Definition 5.12, either the trace is arithmetic and infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x, or it is a truth trace $(A_i)_{i\in\omega}$ and infinitely principal for any rule of the form L_{T*} or L_{F*}, where $* \in \{\wedge, \vee, \forall, \exists, T, F\}$.

By the proof of Theorem 5.11, we know the trace cannot be a progressing arithmetic trace, since this would lead to a contradiction. Thus, the progressing trace is a truth trace $(A_i)_{i\in\omega}$ that is infinitely principal for any rule L_{T*} or L_{F*}. Notice that $A_i \in \Gamma_{n_0+i}$ for any $n_0 \leqslant i$. We construct along the branch a sequence of ordinals $(\alpha_i)_{i\in\omega}$ such that $\alpha_{i+1} \leqslant \alpha_i$ and each α_i is the *least* ordinal satisfying $\mathfrak{K}^{\alpha_i} \models \Gamma_i$ but $\mathfrak{K}^{\alpha_i} \not\models \Delta_i$.

We now show that, whenever $\Gamma_i \Rightarrow \Delta_i$ is the conclusion of a L_{T*} or L_{F*} rule with principal formula A_i then $\alpha_{i+1} < \alpha_i$. Thus, we show that for $A_i = Tr$ or $A_i = Fr$, where r is some term, whenever $\mathfrak{K}^\alpha \models A_i$ then there exists some $\beta < \alpha$ such that $\mathfrak{K}^\beta \models A_{i+1}$. We proceed by doing a case distinction on the left truth and falsity rules.

- $\mathrm{L}_{T\wedge}$: Then $A_i = Tr$ for $r = (s \dot\wedge t)$ and $\mathfrak{K}^\alpha \models Tr$. This means that the value of r is in the α-th approximation of the model, and consequently the value of $s \dot\wedge t$ is as well. By the positive inductive definition of $\mathfrak{K}^\alpha$, there is a $\beta < \alpha$ such that the values of the terms s and t are in the truth set of the β-th approximation of the model, so $\mathfrak{K}^\beta \models Ts, Tt$.

- $\mathrm{L}_{T\forall}$: Then $A_i = Tr$ for $r = \dot\forall x.A$ and $\mathfrak{K}^\alpha \models Tr$. This means that the value of r is in the α-th approximation of the model, and consequently the value of $\dot\forall x.A$ is as well. By the positive inductive definition of $\mathfrak{K}^\alpha$, there is a $\beta < \alpha$ such that the value of the term $A[t/x]$ is in the truth set of the β-th approximation of the model for any $t \in \omega$, so $\mathfrak{K}^\beta \models TA[t/x]$ for any $t \in \omega$.

- L_{TF}: Then $A_i = Tr$ for $r = \dot{F}s$ and $\mathfrak{K}^\alpha \models Tr$. This means that the value of r is in the α-th approximation of the model, and consequently the value of $\dot{F}s$ is as well. By the positive inductive definition of $\mathfrak{K}^\alpha$, there is a $\beta < \alpha$ such that the value of the term s is in the falsity set of the β-th approximation of the model, so $\mathfrak{K}^\beta \models \dot{F}s$.

The remaining cases are similar or trivial. This means that $\alpha_{i+1} < \alpha_i$ for infinitely many i, since infinitely many formulae in $(A_i)_{i\in\omega}$ are principal for the L_{T*} or L_{F*} rules, contradicting the wellfoundedness of the ordinals.

So we again end up with a contradiction, and therefore $\mathfrak{K} \models \Gamma \Rightarrow \Delta$ has to hold for any provable sequent $\Gamma \Rightarrow \Delta$ in Π. $\square$

5.2.3 Proof and Progressivity in CνKF

The progressivity condition defining the cyclic proof system CμKF allows the truth and falsity predicates to be modelled as least fixed points. In the following subsection we will introduce CνKF as being, in a certain sense, dual to CμKF, being able to model the truth and falsity predicates as greatest fixed points. It will roughly be obtained by an alternative progressivity condition that requires truth traces to be infinitely often principal for the right-hand rules for T or F.

Definition 5.16 (Progressing trace in CνKF)**.** A trace in CνKF is *progressing* if it is either an arithmetic trace as in Definition 3.5 and is infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x; or it is a truth trace as in Definition 5.5 and infinitely often principal for any rule of the form R_{T*} or R_{F*}, where $* \in \{\wedge, \vee, \forall, \exists, T, F\}$.

Definition 5.17 (CνKF proof)**.** A pre-proof in CνKF Π is a *proof* in CνKF if along every infinite branch $\Gamma_i \Rightarrow \Delta_i$ there exists some progressing trace along the branch.

The following is a proof in CνKF, and is essentially dual to the proof in Figure 5.2:

$$\textsc{Cut}\ \dfrac{\textsc{Wk}\ \dfrac{\textsc{Ax}\ \dfrac{}{A \Rightarrow A}\ (A \in \Delta)}{\Gamma, A \Rightarrow \Delta, Tz, z = \underset{\cdot}{T}s} \qquad \mathrm{R}_{TT}\ \dfrac{\textsc{Subst}\ \dfrac{\Gamma, A \Rightarrow \Delta, Tz(\star)}{\Gamma, A \Rightarrow \Delta, Ts}}{\Gamma, z = \underset{\cdot}{T}s, A \Rightarrow \Delta, Tz}}{\Gamma, A \Rightarrow \Delta, Tz(\star)}$$

Figure 5.11: Example proof in CνKF

Again, while this is a proof in CμKF, it does not satisfy the conditions of being a proof in CKF. From the progressing trace definitions we have given it should be apparent that CKF is a subsystem of both CμKF and CKF, since any proof in CKF will be a proof in both of the latter systems, but not every proof in these systems will be a proof in CKF.

Embedding νKF into CνKF

Much like in the case for CμKF, we can embed νKF into its cyclic version by showing that the inductive maximality axiom can be subsumed by a cyclic proof with the progressing trace condition we have set as a requirement for CνKF. Therefore any statement that can be derived by an inductive νKF proof can be derived by a proof in CνKF.

Theorem 5.18. $\nu\mathrm{KF} \subseteq \mathrm{C}\nu\mathrm{KF}$

Proof. We show the maximality axiom can be derived by a proof in CνKF. This follows a similar structure to the proof of Theorem 5.14, with the prooftree beginning as follows:

$$
\mathrm{R}_{\to}\,\dfrac{\mathrm{R}_{\forall}\,\dfrac{\mathrm{R}_{\wedge}\,\dfrac{\mathrm{R}_{\to}\,\dfrac{\begin{matrix}\vdots\\ N, P(z) \Rightarrow Tz(\star)\end{matrix}}{N \Rightarrow P(z) \to Tz} \qquad \mathrm{R}_{\to}\,\dfrac{\begin{matrix}\vdots\\ N, Q(z) \Rightarrow Fz(\blacktriangle)\end{matrix}}{N \Rightarrow Q(z) \to Fz}}{N \Rightarrow (P(z) \to Tz) \wedge (Q(z) \to Fz)}}{N \Rightarrow \forall x.((P(x) \to Tx) \wedge (Q(x) \to Fx))}}{\Rightarrow N \to \forall x.((P(x) \to Tx) \wedge (Q(x) \to Fx))}
$$

Figure 5.12: The beginning of the derivation.

The inner nodes marked with $(\star)$ and $(\blacktriangle)$ denote that these nodes are companion nodes for leaves in the continuation of the proof marked with the same symbol. Like in the proof of Theorem 5.14, z must be a sentence since it occurs in Tz, Fz and $P(z), Q(z)$, so we take $\bigvee z$ to be the same disjunction as before, where we know that it is possible to derive $\mathrm{Sent}_{\mathcal{L}^+_{T,F}} \Rightarrow \bigvee z$ through a proof in CPA.

The remainder of the proof goes through in much of the same was as the proof of Theorem 5.14; we first apply a $\textsc{Cut}$-rule, and then the $\mathrm{L}_\vee$-rule (multiple times) to decompose the disjunction $\bigvee z$:

$$
\textsc{Cut}\,\dfrac{\begin{matrix}\vdots\\ \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z) \Rightarrow \bigvee z\end{matrix} \qquad \mathrm{L}_\vee, \mathrm{L}_\vee, \ldots\,\dfrac{\mathrm{L}_\exists, \mathrm{L}_\exists\,\dfrac{\begin{matrix}\vdots\\ z = (s \dot{=} t), N, P(z) \Rightarrow Tz\end{matrix}}{\exists x.\exists y.(z = (x \dot{=} y)), N, P(z) \Rightarrow Tz} \quad \mathrm{L}_\exists, \mathrm{L}_\exists\,\dfrac{\begin{matrix}\vdots\\ z = (s \dot{\wedge} t), N, P(z) \Rightarrow Tz\end{matrix}}{\exists x.\exists y.(z = (x \dot{\wedge} y)), N, P(z) \Rightarrow Tz} \quad \begin{matrix}\vdots\\ \text{other disjuncts of } z\\ \vdots\\ \vdots\end{matrix}}{\bigvee z, N, P(z) \Rightarrow Tz}}{N, P(z) \Rightarrow Tz(\star)}
$$

Figure 5.13: Applying $\textsc{Cut}$ and $\mathrm{L}_\vee$

Likewise along the F-trace on the right-hand branch of Figure 5.12. We will now look at one example of how a branch may develop further.

$$
\begin{prooftree}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$\mathrm{Sent}_{\mathcal{L}^+_{T,F}}, \bar{N}, N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}, P(z), Q(z)$}
\LeftLabel{R$_\vee$}
\UnaryInfC{$\mathrm{Sent}_{\mathcal{L}^+_{T,F}}, \bar{N}, N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}, P(z) \vee Q(z)$}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$\mathrm{Sent}_{\mathcal{L}^+_{T,F}}, \bar{N}, NnP(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}$}
\LeftLabel{L$_\to$}
\BinaryInfC{$(P(z) \vee Q(z)) \to \mathrm{Sent}_{\mathcal{L}^+_{T,F}}, \bar{N}, N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z)$}
\LeftLabel{L$_\forall$}
\UnaryInfC{$\forall x.((P(x) \vee Q(x)) \to \mathrm{Sent}_{\mathcal{L}^+_{T,F}}), \bar{N}, N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z)$}
\LeftLabel{L$_\vee$}
\UnaryInfC{$\bar{N}, N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(z)$}
\LeftLabel{L$_=$}
\UnaryInfC{$\bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz, \mathrm{Sent}_{\mathcal{L}^+_{T,F}}(s \underset{\cdot}{\vee} t)$}
\AxiomC{(cont'd in Figure 5.15)}
\LeftLabel{L$_\to$}
\BinaryInfC{$(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(s \underset{\cdot}{\vee} t) \to (P(s \underset{\cdot}{\vee} t) \to (P(s) \vee P(t)))), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\LeftLabel{L$_\forall$, L$_\forall$}
\UnaryInfC{$\forall x.\forall y.(\mathrm{Sent}_{\mathcal{L}^+_{T,F}}(x \underset{\cdot}{\vee} y) \to (P(x \underset{\cdot}{\vee} y) \to (P(x) \vee P(y)))), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\LeftLabel{L$_\wedge$}
\UnaryInfC{$z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\end{prooftree}
$$

Figure 5.14: $z = (s \underset{\cdot}{\vee} t)$ (1)

$$
\begin{prooftree}
\AxiomC{$N, P(z) \Rightarrow Tz(\star)$}
\LeftLabel{Subst}
\UnaryInfC{$N, P(s), P(z) \Rightarrow Ts$}
\LeftLabel{Wk}
\UnaryInfC{$P(s), \bar{N}, N, P(z) \Rightarrow Ts, Tt$}
\LeftLabel{R$_{T\vee}$}
\UnaryInfC{$P(s), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\AxiomC{$N, P(z) \Rightarrow Tz(\star)$}
\LeftLabel{Subst}
\UnaryInfC{$N, P(t), P(z) \Rightarrow Tt$}
\LeftLabel{Wk}
\UnaryInfC{$P(t), \bar{N}, N, P(z) \Rightarrow Ts, Tt$}
\LeftLabel{R$_{T\vee}$}
\UnaryInfC{$P(t), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\LeftLabel{L$_\vee$}
\BinaryInfC{$P(s) \vee P(t), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\AxiomC{}
\LeftLabel{Ax}
\UnaryInfC{$\bar{N}, N, P(z) \Rightarrow Tz, P(z)$}
\LeftLabel{L$_=$}
\UnaryInfC{$\bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz, P(s \underset{\cdot}{\vee} t)$}
\LeftLabel{L$_\to$}
\BinaryInfC{$P(s \underset{\cdot}{\vee} t) \to (P(s) \vee P(t)), \bar{N}, z = (s \underset{\cdot}{\vee} t), N, P(z) \Rightarrow Tz$}
\end{prooftree}
$$

Figure 5.15: $z = (s \underset{\cdot}{\vee} t)$ (2)

The nodes marked with ($\star$) are buds of the node marked with the same symbol in Figure 5.12. It is easily verified that in every other branch not shown here, the derivation will be similar (but dual) to some branch in Theorem 5.14. Thus any occurrence of the maximality scheme in νKF can be derived in CνKF. In any infinite branch that occurs in such a derivation, we will pass through some rule of the form R_{T*} or R_{F*} infinitely often. This means that any infinite branch has a progressing trace, and therefore the derivation is a proof in CνKF.

□

Soundness of CνKF

We can show that the cyclic system CνKF is sound with the greatest fixed point model as constructed in Section 4.2.2. The proof for the soundness of CνKF follows the same structure as in the case for CμKF. However, first we introduce a new lemma and corollary, nearly identical to Lemma 5.9 and Corollary 5.10 respectively, but take into account the greatest fixed point model:

Lemma 5.19 (Reflecting falsity)**.** *Let $\mathfrak{G}'$ be any fixed point model of the greatest fixed point model construction as in Section 4.2.2. Given an inference step, where* r *is one of the inference rules for* CνKF*:*

$$\mathsf{r}\,\frac{\Gamma_0 \Rightarrow \Delta_0 \qquad \cdots \qquad \Gamma_n \Rightarrow \Delta_n}{\Gamma \Rightarrow \Delta}$$

If $\mathfrak{G}', \rho \models \bigwedge \Gamma$ and $\mathfrak{G}', \rho \not\models \bigvee \Delta$ for some environment $\rho : FV(\Gamma) \cup FV(\Delta) \to \omega$ interpreting the free variables of any formula $A \in \Gamma \cup \Delta$ as natural numbers then there is $\rho' \leqslant \rho$ and $0 \leqslant i \leqslant n$ such that $\mathfrak{G}', \rho' \models \bigwedge \Gamma_i$ and $\mathfrak{G}', \rho' \not\models \bigvee \Delta_i$.

Proof. The proof is dual to the proof of Lemma 5.9; the only change we make is that we replace $\mathfrak{K}'$ with $\mathfrak{G}'$. □

Corollary 5.20. *Let* d *be a be a pre-proof with endsequent $\Gamma \Rightarrow \Delta$ and let $\mathfrak{G}'$ be any fixed point model of the Kripke construction. If $\mathfrak{G}', \rho \not\models \Gamma \Rightarrow \Delta$ then we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ through* d *and a family of environments $(\rho_i)_{i\in\omega}$, with $\rho_{i+1} \leqslant \rho_i$ such that $\mathfrak{G}', \rho_i \not\models \Gamma_i \Rightarrow \Delta_i$ for all $i \in \omega$.*

This allows us to give the soundness result for CνKF, which is in some sense dual to the soundness result for CμKF.

Theorem 5.21 (Soundness)**.** *If* CνKF $\vdash \Gamma \Rightarrow \Delta$ *then* $\mathfrak{G} \models \Gamma \Rightarrow \Delta$.

Proof. Let Π be a CνKF proof of $\Gamma \Rightarrow \Delta$ and suppose towards contradiction that $\mathfrak{G} \not\models \Gamma \Rightarrow \Delta$. By Corollary 5.20, we can construct an infinite branch $(\Gamma_i \Rightarrow \Delta_i)_{i\in\omega}$ and a corresponding family of evironments $(\rho_i)_{i\in\omega}$ falsifying each sequent $\Gamma_i \Rightarrow \Delta_i$. By the progressivity condition of Π we know that the infinite branch has a progressing trace. This means that, by Definition 5.16, either the trace is arithmetic and infinitely often principal for the $\textsc{Case}_x$-rule for the same free variable x, or it is a truth trace $(A_i)_{i\in\omega}$ and infinitely principal for any rule of the form R_{T*} or R_{F*}, where $* \in \{\wedge, \vee, \forall, \exists, T, F\}$.

By the proof of Theorem 5.11, we know the trace cannot be a progressing arithmetic trace, since this would lead to a contradiction. Thus, the progressing trace is a truth trace $(A_i)_{i\in\omega}$ that is infinitely principal for any rule R_{T*} or R_{F*}. Notice that $A_i \in \Delta_{n_0+i}$ for any $n_0 \leqslant i$. We construct along the branch a sequence of ordinals $(\alpha_i)_{i\in\omega}$ such that $\alpha_{i+1} \leqslant \alpha_i$ and each α_i is the *least* ordinal satisfying $\mathfrak{G}^{\alpha_i} \models \Gamma_i$ but $\mathfrak{G}^{\alpha_i} \not\models \Delta_i$.

We now show that, whenever $\Gamma_i \Rightarrow \Delta_i$ is the conclusion of a R_{T*} or R_{F*} rule with principal formula A_i then $\alpha_{i+1} < \alpha_i$. Thus, we show that for $A_i = Tr$ or $A_i = Fr$, where r is some term, whenever $\mathfrak{G}^{\alpha} \not\models A_i$ then there exists some $\beta < \alpha$ such that $\mathfrak{G}^{\beta} \not\models A_{i+1}$. We proceed by doing a case distinction on the right truth and falsity rules.

- $\mathrm{R}_{T\wedge}$: Then $A_i = Tr$ for $r = (s \dot{\wedge} t)$ and $\mathfrak{G}^{\alpha} \not\models Tr$. This means that the value of r is not in the α-th approximation of the model, and consequently the value of $s \dot{\wedge} t$ is not either. By the positive inductive definition of $\mathfrak{G}^{\alpha}$, there is a $\beta < \alpha$ such that the values of the terms s and t are not in the β-th approximation of the model, so $\mathfrak{G}^{\beta} \not\models Ts, Tt$.

- $\mathrm{R}_{T\forall}$: Then $A_i = Tr$ for $r = (\underset{\cdot}{\forall} x.A$ and $\mathfrak{G}^\alpha \not\models Tr$. This means that the value of r is in the α-th approximation of the model, and consequently the value of $\underset{\cdot}{\forall} x.A$ is not either. By the positive inductive definition of $\mathfrak{G}^\alpha$, there is a $\beta < \alpha$ such that the value of the term $A[t/x]$ is not in the β-th approximation of the model for all $t \in \omega$, so $\mathfrak{G}^\beta \not\models TA[t/x]$ for all $t \in \omega$.

- R_{FT}: Then $A_i = Fr$ for $r = \underset{\cdot}{T}s$ and $\mathfrak{G}^\alpha \not\models Fr$. This means that the value of r is in the α-th approximation of the model, and consequently the value of $\underset{\cdot}{F}s$ is not either. By the positive inductive definition of $\mathfrak{G}^\alpha$, there is a $\beta < \alpha$ such that the value of the term s is not in the β-th approximation of the model, so $\mathfrak{G}^\beta \not\models \underset{\cdot}{F}s$.

The remaining cases are similar or trivial. This means that $\alpha_{i+1} < \alpha_i$ for infinitely many i, since infinitely many formulae in $(A_i)_{n_0 \leqslant i \in \omega}$ are principal for the R_{T*} or R_{F*} rules, contradicting the wellfoundedness of the ordinals.

So we again end up with a contradiction, and therefore $\mathfrak{G} \models \Gamma \Rightarrow \Delta$ has to hold for any provable sequent $\Gamma \Rightarrow \Delta$ in Π.

$\square$

Chapter 6

Discussion

In this thesis we have introduced cyclic proof systems for the truth theories KF and μKF, as well as for the theory νKF. This last theory is not discussed in previous literature on theories of truth, but its proof-theoretic analysis provides an interesting extension to the development of the cyclic proof system CKF.

The research outcomes of this thesis contribute to bridging the gap between the realm of cyclic proofs and theories of truth, paving the way to future investigations on the topic. Both thesis contributions and potential research directions are discussed in more detail in the next sections.

6.1 Modularity and arbitrary fixed points in cyclic proofs

The cyclic proof systems CμKF and CνKF presented in this thesis have modular progressivity conditions, meaning that their progressivity conditions rely on a disjunction of different progressing trace definitions. Furthermore, we have shown how to construct the system CKF as a simple fragment of CμKF and CνKF. In other words we have demonstrated that, in certain settings, cyclic proofs can be used to model theories of arbitrary fixed points (not just the least and the greatest ones) such as KF. The inductive theories of truth explored in this thesis each have PA as their base theory, so their cyclic formulations have CPA underlying them. This is what a part of, or in the case of KF what all of, the cyclic reasoning is built on. This also is what allows for the unique presentation of progressivity in our cyclic proof systems, as either being modular or allowing for some arbitrary fixed points. Therefore our systems offer a new perspective on progressivity in the setting of cyclic and ill-founded proofs.

We have shown an embedding of inductive systems into cyclic systems for each of the theories of truth considered. What is more, we have shown that all three cyclic systems are sound with respect to their intended semantic models. However, while we have developed cyclic systems for theories of truth that are sound and embed their inductive counterparts, we still do not know if cyclic and inductive theories of truth prove the same theorems: in other words, it might be the case that our cyclic systems are too permissive in what they can prove in comparison to their inductive counterparts. In the case of CμKF, one could answer this question by exploiting the proof-theoretic equivalence between μKF and ID_1, demonstrated by Burgess in [Bur14] This equivalence should be easily extendable to the cyclic setting. This means that the equivalence CμKF=μKF would be an immediate consequence of the equivalence ID_1=CID_1, where the latter system is the cyclic version of ID_1. However, the equivalence of ID_1 to CID_1 is currently still an open problem (cf. [DM23] for more on cyclic proofs for inductive definitions). Regardless, further analysis of this question is very relevant, since ideally we would like our cyclic and inductive systems to be proof-theoretically equivalent.

6.2 Proof theory for theories of truth

In this thesis we have also strengthened our understanding of the connection between the semantic constructions of the truth predicate and the corresponding syntactic theories of truth, by showing how semantic properties are represented in the setting of cyclic proofs. It would be interesting to further examine this by extending the cyclic proof-theoretic analysis presented in this thesis to other truth theories. Good candidates for this would be the theories PKF and VF, introduced in Section 2.2.4, which would likely allow for similar cyclic systems as KF, since they have roughly the same semantic foundations. Another option is how to proof-theoretically present truth theories that do not use fixed point semantics; for instance a theory like FS, which instead uses revision semantics (cf. [Hal14, p. 145-181]). Furthermore, it might be attractive to develop a cyclic or ill-founded proof system for some of the disquotational truth theories, since these consist of some very simple axioms that would potentially be easy to implement in a cyclic sequent calculus system with progressivity conditions on the truth trace. Despite the fact that disquotational theories are not considered to be very strong, it can be shown that PUTB proves all the same arithmetical sentences as KF [Hal14, p. 268]. Therefore, it would be interesting to see if we can develop a cyclic system for PUTB that is equivalent to CKF in the same way.

Finally, another way to expand on this research is by exploring the relation between cyclic proofs and tableaux or games, seen as syntactic presentations of semantic truth theories (see [Mea15] / [Wel09]). There is some evidence of the interaction between cyclic proofs and games that could support this direction, cf. the cyclic proof-theoretic representation of game logic in [EHK$^+$19]. Looking into this further, and questioning how games connect to cyclic or ill-founded proofs, especially in the context of truth theories, would potentially be valuable in answering some of the open questions above, concerning how we syntactically represent and understand truth.

Bibliography

[AEL22] Bahareh Afshari, S. Enqvist, and Graham E. Leigh. Cyclic proofs for the first-order μ-calculus. *Logic Journal Of The IGPL, 2022*, 2022.

[AL16] Bahareh Afshari and Graham E. Leigh. Finitary proof systems for kozen's μ, 2016.

[ALMT23] Bahareh Afshari, Graham E. Leigh, and Guillermo Menéndez Turata. A Cyclic Proof System for Full Computation Tree Logic. In Bartek Klin and Elaine Pimentel, editors, *31st EACSL Annual Conference on Computer Science Logic (CSL 2023)*, volume 252 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:19, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[BDP11] James Brotherston, Dino Distefano, and Rasmus Lerchedahl Petersen. Automated cyclic entailment proofs in separation logic. In Nikolaj Bjørner and Viorica Sofronie-Stokkermans, editors, *Automated Deduction – CADE-23*, pages 131–146, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[BGP12] James Brotherston, Nikos Gorogiannis, and Rasmus L. Petersen. A generic cyclic theorem prover. In *APLAS*, volume 7705 of *Lecture Notes in Computer Science*, pages 350–367. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[Bro05] James Brotherston. Cyclic proofs for first-order logic with inductive definitions. In *Automated Reasoning with Analytic Tableaux and Related Methods*, volume 3702 of *Lecture Notes in Computer Science*, pages 78–92. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[Bur14] John P Burgess. Friedman and the axiomatization of kripke's theory of truth. *Foundational adventures: Essays in honor of Harvey M. Friedman*, pages 125–148, 2014.

[Can89] Andrea Cantini. Notes on formal theories of truth. *Mathematical Logic Quarterly*, 35(2):97–130, 1989.

[Can90] Andrea Cantini. A theory of formal truth arithmetically equivalent to id 1. *The Journal of symbolic logic*, 55(1):244–259, 1990.

[DM23] Anupam Das and Lukas Melgaard. Cyclic Proofs for Arithmetical Inductive Definitions. In Marco Gaboardi and Femke van Raamsdonk, editors, *8th International Conference on Formal Structures for Computation and Deduction (FSCD 2023)*, volume 260 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:18, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[EHK+19] Sebastian Enqvist, Helle Hvid Hansen, Clemens Kupke, Johannes Marti, and Yde Venema. Completeness for game logic. In *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–13, 2019.

[Fef91] Solomon Feferman. Reflecting on incompleteness. *The Journal of symbolic logic*, 56(1):1–49, 1991.

[Fie94] Hartry Field. Disquotational truth and factually defective discourse. *The Philosophical Review*, 103(3):405–452, 1994.

[Hal01] Volker Halbach. Disquotational truth and analyticity. *The Journal of symbolic logic*, 66(4):1959–1973, 2001.

[Hal14] Volker Halbach. *Axiomatic theories of truth.* Cambridge University Press, New York, USA, revised edition, 2014.

[HH06] Volker Halbach and Leon Horsten. Axiomatizing kripke's theory of truth. *The Journal of symbolic logic*, 71(2):677–712, 2006.

[HL22] Volker Halbach and Graham E. Leigh. Axiomatic Theories of Truth. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Spring 2022 edition, 2022.

[Hor11] Leon Horsten. *The Tarskian turn : deflationism and axiomatic truth.* 2011.

[Ken20] Juliette Kennedy. Kurt Gödel. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Winter 2020 edition, 2020.

[Kri75] Saul Kripke. Outline of a theory of truth. *The Journal of philosophy*, 72(19):690–716, 1975.

[KS16] Ioannis Kokkinis and Thomas Studer. *Cyclic Proofs for Linear Temporal Logic*, pages 171–192. De Gruyter, Berlin, Boston, 2016.

[Lei13] Graham E Leigh. Conservativity for theories of compositional truth via cut elimination. *arXiv.org*, 2013.

[Lei16] Graham E. Leigh. Reflecting on truth. *IfCoLog Journal Of Logics And Their Applications, 2016, Vol. 3, Iss. 4, pp. 557-.594*, 3(4):557–594, 2016.

[McG90] Vann McGee. *Truth, vagueness, and paradox : an essay on the logic of truth.* Hackett, Indianapolis, Ind., 1990.

[Mea15] Toby Meadows. Infinitary tableau for semantic truth. *The review of symbolic logic*, 8(2):207–235, 2015.

[Mos74] Yiannis Nicholas Moschovakis. *Elementary Induction on Abstract Structures.* Dover Publications, Mineola, N.Y., 1974.

[Mur20] Roman Murawski. Proof vs truth in mathematics. *Studia Humana (Rzeszów)*, 9(3):10–18, 2020.

[Rei86] William N. Reinhardt. Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of philosophical logic*, 15(2):219–251, 1986.

[Sim17] Alex Simpson. Cyclic arithmetic is equivalent to peano arithmetic. In Javier Esparza and Andrzej S. Murawski, editors, *Foundations of Software Science and Computation Structures*, pages 283–300, Berlin, Heidelberg, 2017. Springer Berlin Heidelberg.

[Tar31] Alfred Tarski. The concept of truth in formalized languages. In A. Tarski, editor, *Logic, Semantics, Metamathematics*, pages 152–278. Oxford University Press, 1931.

[Tar55] Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285 – 309, 1955.

[Tar69] Alfred Tarski. Truth and proof. *Scientific American*, 220(6):63–77, 1969.

[TB20] Gadi Tellez and James Brotherston. Automatically verifying temporal properties of pointer programs with cyclic proof. *Journal of automated reasoning*, 64(3):555–578, 2020.

[TU22] Takeshi Tsukada and Hiroshi Unno. Software model-checking as cyclic-proof search. *Proc. ACM Program. Lang.*, 6(POPL), jan 2022.

[Weh23] Dominik Wehr. Representation matters in cyclic proof theory, 2023.

[Wel09] P. D. Welch. Games for truth. *The bulletin of symbolic logic*, 15(4):410–427, 2009.