# Categories with New Foundations 



## Adam Vincent Lewicki

## University of Cambridge

Department of Pure Mathematics and Mathematical Statistics

Clare Hall


March 2020

## Declaration

This dissertation is the result of my own work (Chapters 1, 2 and 5; Sections 3.3, 4.3.1 - 4.5, and Section 4.6 excluding Theorem 4.48) and collaboration with Thomas Forster and Alice Vidrine (Sections 3.1-3.2, 3.4-3.6, 4.1-4.3.0 up to and including Theorem 4.19, and Theorem 4.48). Chapter 1, and Sections 2.1, 4.1, 4.4.1, 4.5 .1 and 5.1, consist mostly of review material, with original results only where indicated. Throughout this work, I have made every effort to employ direct citations (i.e. 'Theorem[citation]') where results reference specific works. Where a result is widely known, I may cite a textbook, or assume the context obviates a specific reference, rather than the original paper. As this thesis studies largely disparate areas of research, NF Set Theory and Category Theory, I have adopted a preference for self-containment, in terms of background and introductory material. Where I chose to include information, rather than refer the reader to review specific text(s), I have indicated those I followed most closely. In doing so, I acknowledge any novelty to be theirs and any error to be my own.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Acknowledgements and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Acknowledgements and specified in the text.

# Categories with New Foundations 

Adam Lewicki

## Summary

While the interaction between set theory and category theory has been studied extensively, the set theories considered have remained almost entirely within the Zermelo family. Quine's New Foundations has received limited attention, despite being the onesorted version of a theory mentioned as a possible foundation for Category Theory by Mac Lane and Eilenberg in their seminal paper on the subject 6].

The lack of attention given to NF is not without justification. The category of NF sets is not cartesian closed and the failure of choice is a theorem of NF [40, 61. But those results should not obscure the aspects of NF that have foundational appeal, nor the value of studying category theory in the context of a universal set.
$\mathcal{N}$, the category of NF sets, provides a closed foundation for category theory:

$$
\mathcal{N} \in \operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\mathcal{N})
$$

Or, iterating enrichment:

$$
\mathcal{N} \in \operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\operatorname{cat}(\mathcal{N})) \in \operatorname{cat}^{3}(\mathcal{N}) \ldots \operatorname{cat}^{n}(\mathcal{N}) \in \operatorname{cat}^{n+1}(\mathcal{N}) \ldots
$$

where $\operatorname{cat}^{n}(\mathcal{N})$ is the internal category of $n$-categories in $\mathcal{N}$.

In addition to higher categories, typically large categories of interest (e.g. Top, Cat, Set, $G r p$ ) are small categories in NF Furthermore, the internalisation of even the largest of these categories displays surprising coherence. Despite the failure of cartesian closure, the internalisation of $\mathcal{N}$ into itself turns out to be a full internal subcategory.

[^0]Further still, while $\mathcal{N}$ also lacks a number of important adjoint relationships beyond cartesian closure, which are taken for granted in Set, it possesses corresponding relative adjunctions. ${ }^{2}$

What are perceived to be advantages and disadvantages of working in $\mathcal{N}$ turn out to be two sides of the same coin. Set theoretic implementations of mathematical objects in NF often reflect the "naive" implementation of Frege more than the familiar implementation in $\mathrm{ZF}(\mathrm{C}) \cdot{ }^{3}$ Similarly, category theoretic properties of $\mathcal{N}$ reflect many of the complexities we encounter (often informally) when working in the (naive) category of all categories.

The informal connection between $C A T$ and $\mathcal{N}$ can be formalized by thinking of sets as $\{T, \perp\}$-enriched categories. Presheaves can be seen as Set-enriched generalizations of powersets [4]. Thus, in turn, powersets are $\{T, \perp\}$-enriched presheaves. Classically, $P:$ Set $\rightarrow$ Set is a monad, while $(\widehat{-}): C a t \rightarrow C A T$ is a relative pseudomonad [8]. In $\mathcal{N}$, on the other hand, both $P: \mathcal{N} \rightarrow \mathcal{N}$ and $(\widehat{-}): \operatorname{cat}(\mathcal{N}) \rightarrow \operatorname{cat}(\mathcal{N})$ form relative algebraic structures. In this way, what is a "weakness" of $\mathcal{N}$ in the context of Set, is a form of "coherence" in the context of Cat. In the same way, the $\mathcal{Y}$-relative adjunction, forming part of the Yoneda Extension diagram (see 4, 65), has a "degenerate" form in $\mathcal{N} .1$ In fact, the $T$-functor, imbricate with categories of stratified sets, can itself be seen as a form of Yoneda Embedding - both arise as units to relative (pseudo)monads.

The relationship between $\mathcal{N}$ and Cat is maintained, even in the case where category theory is developed in NF (i.e. cat $(\mathcal{N})$ ). Fam, for example, is an internal lax idempotent relative pseudomonad, in the sense of [8]. The relationship between powersets and presheaves is also maintained and, in some ways, strengthened. While the universe of sets is closed under $P$ in both $\mathrm{ZF}(\mathrm{C})$ and NF, the same is not true of $(\hat{-})$ in the classical theory of categories, developed over $\mathrm{ZF}(\mathrm{C})$. Whereas one can continually iterate $P$, the presheaf functor, $(\widehat{-})$, is only defined for locally small categories. In $\mathcal{N}$, the presheaf

[^1]functor remains a relative algebraic structure (due to its inhomogeneity), but $\operatorname{cat}(\mathcal{N})$ is closed under presheaves (i.e. $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$ implies $\widehat{\mathbb{C}} \in \operatorname{cat}(\mathcal{N})$ ).

The present research expands upon the author's work in [14] (much of which is contained in Chapter 3). It is not intended to "advocate" for the use of NF as a practical foundation for category theory. Instead, the work presents a broad survey of the interaction between the set theory and category theory of NF, examining the relationship in both directions. The abstract structure, of which both type restriction (in $\mathcal{N}$ ) and size restriction in $C A T$ are specific cases, appears to be the study of relative algebra. In a number of cases, the existence of a relative algebraic structure in $\mathcal{N}$ can be proven more generally for a class of relative adjoints, (pseudo)monads, etc. Thus, where it seems appropriate to do so, this thesis seeks to contribute to the broader study of relative algebra. ${ }^{5}$

## ***** Overview of Chapters *****

Chapter 1 motivates and provides an introduction to the multi-sorted theory of TST and the single-sorted theory of NF [45].] Outside of some basic results regarding category theory in TST(I), this chapter is comprised primarily of review material. Truly introductory accounts of NF are lacking, in general, so our instinct is toward self-containment. ${ }^{7}$ This chapter is tailored to readers who are less familiar with NF. That said, even those who are well acquainted with NF may want to note examples and basic results relating to the category theory of NF. In particular, the introduction of lateral functions and lateral functors, in ML, will be important for understanding the category theory of $\mathrm{NF}^{8}$

Central to the development of category theory in NF is the exchange of classical adjoint relationships for relative adjoints. In Chapter 2, we prove some results relating to the interaction between relative and classical adjoint functors, including a form of relative

[^2]monadicity theorem. These more general results will be required for the study of monadic and (relative) comonadic presentations of internal presheaves in $\mathcal{N}$. A feature of relative adjoints in $\mathcal{N}$ is their recurrence in the form of a canonical pair satisfying a stronger form of adjoint symmetry. We refer to these as symmetric lifts. Proposition 2.17 shows that Yoneda Extensions provide an example of symmetric lifts, outside the context of NF, ${ }^{9}$

Much of Chapter 3 is comprised of collaborative work with Thomas Forster and Alice Vidrine $\sqrt{10}$ This chapter studies the general properties of a category of stratified sets, $\mathcal{N}$. The primary focus is NF, but we also consider KF and some relevant extensions [13. We show $\mathcal{N}$ to be a regular category, admitting a symmetric lift, in place of the classical exponential adjunction, $A \times(-) \dashv(-)^{A}$. We prove an analogue to the Fundamental Theorem of Topos Theory for an NF-topos (Definition 3.18), exchanging classical for modified dependent products. The final section employs ideas from Algebraic Set Theory, to probe at issues of size in NF. Extending the folklore definition of "small sets" in NF - the strongly cantorian sets - to an appropriate category theoretic (i.e. fibre-wise) definition of smallness requires extending the axioms of NF to include $S C U,{ }^{T 1}$ While the motivation for introducing NF $+S C U$ was category theoretic, the resulting extension turns out to be of set theoretic interest, as well ${ }^{12}$ Section 3.3 is the one part of this chapter which is not considered in [14]. Here we study the nature of universal properties in the presence of unrestricted (stratified) comprehension. The existence of large sets permits the formation of a set of structures, among which an initial/terminal subset is universal.

Chapter 4 takes a dual approach to Chapter 3. Rather than studying the category of NF

[^3]sets, we look at the broader implications of taking NF as a base theory of sets in/over which to develop category theory. That is, we develop category theory internal to $\mathcal{N} \cdot{ }^{13}$ The most obvious appeal is the straightforward construction of both $\mathcal{N}$ and $C a t$ as internal categories, $\tilde{\mathcal{N}}$ and $\operatorname{cat}(\mathcal{N})$, from which we obtain the paradigmatic result that $\mathcal{N}$ internalizes itself. However, we have not so much obviated size restrictions as replaced them with stratification/homogeneity restrictions. While this complicates the internal theory of categories, we are, nevertheless, able to prove a stratified Yoneda Lemma.

The NF-Yoneda Lemma is a first step toward a broader study of the presentation of presheaves as algebras and relative coalgebras. Here we apply much of what is developed in Chapter 2.

Regarding $\operatorname{cat}(\mathcal{N})$ as a 2-category enables us to further develop these ideas in the context of relative KZ-Doctrines. The internalisation of Fam in $\mathcal{N}$ forms an internal lax idempotent relative pseudomonad, in the sense of [8]. This leads us to ask more general questions about KZ-(pseudo)algebras in the relative case - in particular, the classical result that KZ-(pseudo)algebras are classified as adjoint to the unit does not appear to extend beyond free relative pseudoalgebras. The final section returns to the study of Fam, now in the context of the family fibration (rather than coproduct completion). The aim is to better understand the role of $\tilde{\mathcal{N}}$, the full internal subcategory of NF sets, and the (stratifed) membership relation $\epsilon_{\mathcal{N}}$ as both the generating $V$-indexed family of $\tilde{\mathcal{N}}$ and object classifier in $\mathcal{N}$. Despite being a full internal subcategory generated by $\epsilon_{\mathcal{N}}$, aspects of $\tilde{\mathcal{N}}$ appear more similar to a universe of types than a universe of sets.

While the interaction between category theory and NF is the primary motivation for Chapter 4, the study of free coproduct completion in $\operatorname{cat}(\mathcal{N})$ motivates a more general examination of relative KZ-Doctrines ${ }^{14]}$ There is significant "empirical" evidence that lax idempotent relative pseudomonads $5^{15}$ are the appropriate generalization of standard KZ-Doctrines to the relative case. Indeed, multiple results in Chapter 4 add to this

[^4]body of evidence. However, the classification theorem for KZ-(pseudo)algebras in the standard case - algebra maps correspond to 1-cells, adjoint to the unit of the monad does not extend to lax idempotent relative pseudomonads. ${ }^{16}$ Nevertheless, one should not conclude from Lemma 4.44 that lax idempotent relative pseudomonads are an incorrect generalization of KZ-Doctrines. It seems more likely that, as with relative monadicity (Theorem 2.33), what is a straightforward classification result in the standard theory of monads, is case-dependent in the relative theory. In other words, the relevant question is what properties, satisfied by the trivial (i.e. identity) relative functor, are required for the classification result to hold in the relative case? Given the potential applications of relative KZ-Doctrines, this strikes the author as an important area for further research.

Chapter 5 pivots from general category theory to $\lambda$-calculus in NF. It is easy to show that $|V \Rightarrow V| \cong|V|$, making NF a tempting model for untyped $\lambda$-calculus.$^{17}$ Scott conjectured that one could implement a multi-relation model (see [58]) of untyped $\lambda$ calculus in a model of NF. But the implication of Scott's conjecture is, in fact, even stronger than it appears. Both Plotkin and Scott have provided a broad class of models (of which multi-relation model is a special case) of untyped $\lambda$-calculus that can be constructed in ZF 44, 58]. The implication of Scott's conjecture for NF, on the other hand, is that any model of NF has a natural interpretation as a model of untyped $\lambda$ calculus - each object in the model is simultaneously a set and a combinator. Chapter 5 proves Scott's conjecture.

Scott further conjectured that one could interpret the sets of NF as finite sequences, in such a way that there was an exact equivalence, $V^{*}=V$, between the collection of all finite sequences and the universe of NF sets. The latter conjecture motivates a reexamination of the standard implementation of finite sequences as nested Quine pairs.

[^5]Surjectivity of Quine pairs implies:

$$
\forall n \in N . V^{n}=V
$$

But the result is not cumulative, in the sense of $V^{*}$. Furthermore, by recursively extending the standard implementation to form streams (i.e. $\omega$-sequences), we obtain the surprising result: $V^{\omega} \neq V$. This motivates the introduction of an alternative implementation of streams, Quine sequences, from which we obtain the identity $V^{\omega}=V$. From this equivalence, we obtain an extension of Scott's conjecture, $V^{*}=V$, to include (possibly) infinite sequences.

The inclusion of $\omega$-sequences also extends the multi-relation model and, therefore, has a non-trivial impact on certain important combinators - the $S$-combinator, in particular - implying a connection between continuity of $S$ and (weak) choice principles. To study this more formally, we introduce pre-combinatory algebras, thinking of sequence formation as a special case of a more general "coding" system. The generalization allows us to study the relationship between the strength of a given coding system and the resulting collection of continuous total functions.

## Acknowledgments

Excessive thanks often serve to advertise excessive accomplishment. In this case, I fear my work does little justice to the contribution of all those who have helped and supported me.

My first, insufficient show of gratitude must be to my parents and family. In everything, I am so grateful to have each one of you. My grandfather, in particular, deserves special recognition. Were it not for his support (financial but, more importantly, emotional) and encouragement, I would never have applied for, nor pursued this course. Not to be outdone, my grandmother also deserves special recognition for walking me through my first piece of schoolwork (at age 6) on mathematical logic $\sqrt{18}^{18}$

I owe significant gratitude to my supervisor, Thomas Forster. His guidance and encouragement has been invaluable to me since my first day at Cambridge. Throughout my time I benefited greatly from my peers in the Category Theory and Logic group, in particular: Phillip K., Zacherie, Sean, Zhen Lin, Filip B., and Enrico. Professors Hyland and Johnstone have been unwavering in their patience, always taking the time to kindly answer my many questions and review my work. I am grateful to have had the opportunity to learn from them. Special thanks to Julia Goedecke for taking me on as a student and for introducing me to category theory. Beyond those I have named, I would like to thank the Department of Mathematics and Clare Hall for providing a home during my time at Cambridge.

Outside of Cambridge, Randall Holmes and Dana Scott have been exceedingly generous in their time and support. Their past work and suggestions have inspired a large portion

[^6]of this thesis. Paul Taylor has also been influential, and generous with his notes and assistance on many topics. Grant Passmore has been both a wonderful friend and teacher. From my time at Davidson, the teaching and friendship of Professors Robert Whitton, John Swallow, Donna Molinek, Paul Studtmann, Mark Foley, Alicia Sparling and Greg Snyder has been of constant value. Finally, Alice Vidrine has been as much a second supervisor as a collaborator. I am a constant admirer of her intellect and creativity, and the present work would hardly have been possible without her. What is more, she has been a constant friend during a process whose challenges are by no means confined to the intellectual.

None of my work would have been worthwhile without the support and company of my friends. I do not wish to name people, for I am certain I will not be able to name all those who deserves thanks. Friends I cannot avoid naming are: Tim, Mali, Sara Page, Nino, Sara, Floor, Pierre, Edouard, Douglas, Nick, Siobhan, Kristen, Marcus, Mac, Amber, Natasha and Grant (again). Luna, I really cannot even attempt to express what you mean to me. It is no overstatement to say that I would not have completed this work without you. And yet, even this would not be among the first hundred ways I would list, in which I have been touched by your kindness and brilliance.

I would like to dedicate this work to the memory of Dr. Robert Whitton. In math and in life, I would know so much less were it not for your teaching.

## Contents

1 Preliminaries ..... 21
1.1 Introduction ..... 21
1.2 An Introduction to TST ..... 25
1.2.1 Category Theory in TST(I) ..... 27
1.2.2 Canonical Embeddings ..... 29
1.3 An Introduction to NF ..... 34
1.3.1 The Natural Numbers ..... 37
1.3.2 A Homogeneous Pairing Function ..... 40
1.3.3 Cantor's Theorem ..... 41
1.3.4 Cantorian and Strongly Cantorian Sets ..... 42
1.3.5 Ordinals and Cardinals in NF ..... 44
1.4 Important Variants of NF ..... 48
1.4.1 KF ..... 49
1.4.2 NFU ..... 51
1.4 .3 ML ..... 52
1.5 Appendix: Typical Ambiguity, from TST to NF ..... 55
2 Relative Algebra ..... 61
2.1 A Generalization of Adjunction ..... 62
2.1.1 Composition of Relative Adjoints ..... 65
2.2 Symmetric Lifts ..... 67
2.2.1 Yoneda Extensions as Symmetric Lifts ..... 68
2.3 Relative (Co)Monads ..... 76
2.3.1 A Canonical (Relative) Monad ..... 81
2.4 Monadicity: The Relative Case ..... 85
3 Categories of Stratified Sets ..... 95
3.1 Defining a Category of Stratified Sets ..... 97
3.1.1 Set Theories of Interest ..... 97
3.1.2 The Basic Category ..... 98
3.1.3 The T Functor ..... 110
3.1.4 A More General Category ..... 114
3.2 Basic Categorical Properties ..... 115
3.2 .1 KF ..... 115
3.2 .2 NF ..... 120
3.3 Universal Properties with Universal Sets ..... 126
3.4 Modified-Cartesian Closure ..... 133
3.5 Local Modified-Cartesian Closure ..... 140
3.6 Smallness Conditions for NF ..... 147
3.6.1 Algebraic Set Theory of NF ..... 147
3.6 .2 Consequences of SCU ..... 154
4 Category Theory in NF ..... 161
4.1 Basic Internal Category Theory ..... 164
4.2 Standard Internal Presentation of Presheaves ..... 170
4.2.1 The Internal Yoneda Lemma ..... 172
4.2.2 Presheaves as Algebras ..... 173
4.2.3 Presheaves as Coalgebras ..... 174
4.3 The Internal Presentation of Presheaves in $\mathcal{N}$ ..... 175
4.3.1 The NF Yoneda Lemma ..... 175
4.3.2 The Presheaf Relative Comonad ..... 176
4.3.3 Algebraic and Relative Coalgebraic Presentations ..... 184
4.4 An Internal Relative KZ-Monad in $\mathcal{N}$ ..... 188
4.4.1 Externalisation and Fam ..... 188
4.4.2 Relative Pseudomonads ..... 191
4.4.3 Internalizing Fam in $\mathcal{N}$ ..... 194
4.5 Cocomplete Objects and Relative KZ-Algebras ..... 200
4.5.1 Relative Pseudoalgebras ..... 200
4.5.2 Are All Relative KZ-algebras Cocomplete? ..... 207
$4.6 \tilde{\mathcal{N}}$ ..... 219
$4.6 .1 \quad \operatorname{Fam}(\tilde{\mathcal{N}})$ ..... 220
4.6.2 $\tilde{\mathcal{N}}$ As A Universe ..... 223
5 NF Lambda Calculus ..... 235
5.1 Scott's Model ..... 239
5.1.1 Distributive and Continuous Functions in NF ..... 249
5.2 A Calculus of Distributive Functions ..... 251
5.2.1 Category of Retracts ..... 254
5.3 A Model of The $\lambda$-Calculus ..... 260
5.3.1 Coding Sequences of Varying Length ..... 261
5.3.2 Implementation of Application and Abstraction ..... 262
5.3.3 Implementation of Combinators ..... 265
5.4 NF With Surjective Sequences ..... 272
5.4.1 Quine Pairs, A Further Look ..... 272
5.4.2 Quine Sequences ..... 278
5.4.3 A "Cumulative Hierarchy" of Sequences ..... 280
5.4.4 Application of Streams ..... 282
5.5 Generalizing Sequences to Codes . . . . . . . . . . . . . . . . . . . . . . 286
5.5.1 Pre-Combinatory Algebra . . . . . . . . . . . . . . . . . . . . . . 292

Bibliography 308

## Chapter 1

## Preliminaries

### 1.1 Introduction

Neither category theory nor set theory share a formal dependency on the other. The two fields are imbricate, both from a foundational perspective and for a "working mathematician" [33]. Set theory provides motivating examples for many abstract categorical structures, but it also plays a more formal role. The category of sets is the most natural mathematical universe within which one can develop a theory of (small) categories. Even for "large" categories one requires an ambient theory of collections - more formally, a chosen category over which one can develop the broader theory of categories [3]. In the other direction, any model of a coherent set theory will form a category. Much of the set theory one develops with the standard membership relation ' $\in$ ' can be developed with generalized elements, in the corresponding category. Standard constructions in Set are often special cases of important structures in CAT, where sets are viewed as discrete categories, enriched over $\{T, \perp\}$.

Broadly speaking, we can categorize the body of research into the relationship between category theory and set theory in the same manner as the examples we have given of what one theory does for the other. There are a number of examples that utilize category
theory to inform the study of set theory ${ }^{\top}$ However, more often than not, the literature works in the opposite direction.

From the standpoint of foundations, almost all mathematical research occurs (although rarely more than "implicitly") within a model of set theory. In many ways, category theory is no exception. As most categories of interest (Top, Grp, etc.) are "large," however, there is a need for a broader theory of classes or, at least, an ambient set that satisfies closure conditions, such as a Grothendieck Universe. Indeed, even the category of small categories, internal to a given model of set theory, is not closed under presheaves in the way Set is closed under powersets (presheaves enriched in $\{\top, \perp\}$ ).

The chosen model of set theory has more "practical" implications for the development of category theory as well $2^{2}$ Universal structures are defined only up to isomorphism, but one typically classifies a canonical universal structure, functorially [34]. Implicit in this is the equipment of categories with "chosen structure." Furthermore, while it can be studied abstractly, the dependence of category theory on a theory of collections (i.e. indexed families and the more general study of fibred categories) implies the relevance of an underlying theory of sets extends beyond (locally) small categories.

In this entire body of research, however, little attention is paid to Quine's New Foundations [45]. This is not entirely unjustified, from a practical standpoint in particular. The only prominent result in the literature is McLarty's proof that cartesian closure fails in any consistent model of NF [40]. Furthermore, NF proves the failure of the axiom of choice [61]. Nevertheless, there are many aspects of $\mathcal{N}$, the category of NF sets, which remain tempting. In particular, size does not preclude the existence of sets in NF. As a result, most categories of interest are small $:^{3}$ Cat, Top, Grp, FamC, etc. Indeed, we can prove the paradigmatic result:

$$
\mathcal{N} \in \operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\mathcal{N})
$$

Not only can $\mathcal{N}$ internalize itself, but it can do so in a coherent manner. Even with

[^7]the failure of cartesian closure, the internal category of NF sets, $\tilde{\mathcal{N}}$, is a full internal subcategory.$^{4}$

The origin of the present work is [14], which studies the general category theoretic properties of stratified set theories (primarily KF and NF). In this paper, we introduced the idea of modified-cartesian closure. While $\mathcal{N}$ does not possess the adjunction $(A \times-) \dashv$ $(-)^{A}$, it possesses a pair of relative adjunctions, which form a stratified analogue of cartesian closure 5 Furthermore, just as cartesian closure is stable under localization in a topos, $\mathcal{N}$ is locally modified-cartesian closed.

While one can think of modified-cartesian closure as "stratified" cartesian closure, the analogy is imperfect. Relative adjunctions possess only half of (standard) adjoint symmetry. But relative structures are not confined to the periphery of category theory: the presheaf monad, for example, is a relative pseudomonad; and, given $F: \mathcal{A} \rightarrow \mathcal{B}$, we obtain the relative $\mathcal{Y}_{\mathcal{A}}$-relative adjoint, $F \mathcal{Y}_{\mathcal{A}} \dashv \mathcal{B}(F-,-)$, familiar from the study of Yoneda Extensions [8, 4]. If we think of sets as $\{\top, \perp\}$-enriched categories (standard (locally small) categories being Set-enriched), we obtain analogous results in $\mathcal{N}$. The powerset functor forms a relative monad, corresponding to the $\{\top, \perp\}$-enriched presheaf (pseudo)monad (Proposition 2.25).


The Yoneda Extension, meanwhile, can be seen as a pasting of $\mathcal{Y}$-relative adjoints, "approximating" $L_{\mathcal{A}}(F)$, in the same sense that modified-cartesian closure approximates the product-exponential adjunction ${ }^{6}$


[^8]In general, this should not be thought of as a strict, formal correspondence. But one of the main goals of the present work is to study not only the category theory of NF, but the way in which the complexities of $\mathcal{N}$ reflect those which have long been known to arise in $C A T$, the "universe" of categories $\sqrt[7]{ }$

Evaluating NF as a foundation for category theory (to be clear, we are not advocating this) also requires a further step. The connections we are drawing between $\mathcal{N}$ and $C A T$ should remain stable under the move from a "generic" universe of categories to $\operatorname{cat}(\mathcal{N})$. This largely turns out to be the case. Fam, for example, turns out to be an internal relative KZ-pseudomonad (Theorem 4.33). 8 In this way, unlike Cat and Set, cat $(\mathcal{N})$ and $\mathcal{N}$ possess similar closure conditions.

Despite the potentially attractive features of $\mathcal{N}$, we should not suppress the undesirable complications of relative algebraic structures, which go beyond asymmetry. Relative (pseudo)algebras often lack free presentations (which is not to say that relative (pseudo)monads do not have free algebras). The relationship between the algebraic presentation of (internal) presheaves and the relative coalgebraic presentation is not as clean as we would like (Theorem 4.24). Nor are we able to prove that the left extension property of relative KZ-pseudomonads classifies relative KZ-pseuodalgebras (specifically, non-free pseudoalgebras), as in the standard case [26]. In addition, NF does not eliminate the need for external categories and functors, including those of significant interest to us ${ }^{9}$

Where there are challenges, we also see interesting avenues for further research. Many of the relative structures we take from $C A T$ could equally exist in a typed setting, such as $T S T$ (typed set theory) ${ }^{10}$ This observation was made (albeit in passing) in the original work of Eilenberg and Mac Lane [6]. Contrary to TST, however, we want to think of the

[^9]"hierarchy" of categories as cumulative rather than disjoint (i.e. a small category should also be a locally small category). Thus, the way in which TST relates to the single sorted theory of NF could inform our understanding of not only $C A T$, but also higher category theory. Just as $\operatorname{cat}(\mathcal{N})$ is closed under presheaves, one can iterate enrichment:
$$
\operatorname{cat}(\mathcal{N}), \operatorname{cat}(\operatorname{cat}(\mathcal{N})), \ldots, \operatorname{cat}^{n}(\mathcal{N}), \ldots
$$
$\operatorname{cat}^{n}(\mathcal{N})$ is the internal $(n+1)$-category of $n$-categories. Lower levels can be embedded in successive levels by adding trivial (i.e. identity) cells, just as sets are embedded into higher levels (in effect, by $\{\cdot\}: X \rightarrow P X$ ) in TST and in the syntax of NF.

### 1.2 An Introduction to TST

The most prominent paradox in naive set theory is Russell's Paradox:

$$
\{x \mid x \notin x\}
$$

Zermelo-style set theories handle this by implementing size restrictions - a class that is not too big is a set. Thus, comprehension becomes separation. As a consequence, one cannot form the universal set:

$$
V=\{x \mid x=x\}
$$

If one could, the axiom of separation would be equivalent to full comprehension.

An alternative way of handling the paradox, also due to Russell, is to focus on the syntactic complexity of the predicates themselves [51]. One of the primary motivations and the most basic intuition for set theory is the study of mathematical objects which are "predicates-in-extension." Intuitively, $\{x \mid \Phi(x)\}$ both defines and is defined by the predicate $\Phi(x)$. Through this lens: the most basic sets should correspond to the most basic predicates.

$$
\emptyset=\{x \mid x \neq x\} \text { and } V=\{x \mid x=x\}
$$

The problematic set $\{x \mid x \notin x\}$ is defined by an intuitively pathological predicate. Therefore, obtaining consistency by restricting the "size" of sets both precludes the existence
of a set corresponding to the most basic predicate (self-identity) and fails to address the "pathological" aspect of $\{x \mid x \notin x\}$, the syntax of the predicate. Russell's solution was to impose grammatical rules: ill-formed predicates do not define sets, just as not all sequences of words convey meaning. The resulting theory was Russell's theory of types 51]. "Typed Set Theory," TST, represents a further evolution of Russell's original idea.

TST is a many-sorted theory, with countably many levels, $T_{0}, T_{1}, \ldots$ and an infinite supply of "typed" variables, at each level. The primitive binary predicates are equality $={ }_{n}$ at each level, and set membership $\epsilon_{n}$ between any two levels $n$ and $n+1$. In this way the objects of level $n+1$ are precisely the sets of objects at level $n$, with level 0 interpreted as the level of "individuals."

A formula in the language of TST is well-formed if each occurrence of ' $x_{n} \in_{n} y_{n+1}$ ' and ' $x_{n}={ }_{n} z_{n}$ ' is typed in the appropriate manner.

Definition 1.1 (TST). TST is defined by the axiom scheme:

## 1. Extensionality at each level.

$$
\forall x_{i+1}, y_{i+1} \cdot\left(x_{i+1}={ }_{i+1} y_{i+1}\right) \Longleftrightarrow\left(\forall z_{i} \cdot z_{i} \in_{i} x_{i+1} \Longleftrightarrow z_{i} \in_{i} y_{i+1}\right)
$$

## 2. Comprehension at each level.

$$
\forall \vec{x} \exists y_{i+1} \forall z_{i} . z_{i} \in y_{i+1} \Longleftrightarrow \phi\left(\vec{x}, z_{i}\right)
$$

where $\phi\left(\vec{x}, z_{i}\right)$ is a well formed formula of TST.

TST admits standard set theoretic constructions, but many are inhomogeneous. Powerobjects and function spaces ${ }^{[1]}$ occur one level above the sets from which they are derived. $T_{0}$, the "individuals," can be thought of as anything one likes: groups, spaces, etc. These are not quite atoms, however. TSTU, TST + urelemente (atoms), admits a collection of atoms at each level.

Definition 1.2. Common extensions of TST:

[^10]1. TSTU, the theory of TST with a collection of atoms at each level.
2. TSTI, the theory of TST with an infinite collection of individuals (i.e. $T_{0}$ is infinite).
3. $\mathbf{T S T}+\mathbf{A C}$, the theory of TST with the axiom of choice ${ }^{12}$
4. TST $+\mathbf{A m b}$, the theory of TST with Ambiguity will be defined later in this section. It is shown to be equiconsistent with NF [62].

TST has a canonical model (even) in Zermelo set theory. Take any set $X$ as the level $M_{0}$ of individuals, then iterate the powerset operation. $\left\{X, P X, P^{2} X, \ldots\right\}$ is a model of TST, where $M_{i} \sim P^{i} X$. We can internalize the construction in an elementary topos $\mathcal{E}$ :

Theorem 1.3. The iterated powerobject chain over an object $C$ in the topos $\mathcal{E}$ forms an internal model of TST. Furthermore, an appropriate colimit taken over this diagram yields a model of Mac Lane Set Theory with "Quine Atoms" (i.e. a set x, such that $x=\{x\}$ ).

### 1.2.1 Category Theory in TST(I)

In their seminal paper introducing category theory, Eilenberg and Mac Lane mentioned the potential that a system like TST could be taken as a foundation [6]. To the author's knowledge, however, no mention of "typed category theory" has appeared in the literature since. While TST formally addresses some basic foundational issues of category theory, most basic properties would require the exchange of adjunctions for the weaker, asymmetric notion of adjunctions that are relative to an external, type-shifting functor. ${ }^{13}$ One would have to reason externally to the theory, regardless.

Remark (Injective Type Shifting). An example of "external" reasoning would be to declare that, for a given set $x_{i}$ :

$$
\left|x_{i}\right| \cong\left|\left\{\left\{z_{i-1}\right\} \mid z_{i-1} \in x_{i}\right\}\right|
$$

[^11]The bijection is obvious: $z_{i-1} \mapsto\left\{z_{i-1}\right\}$, but the graph of this function is not a set, as $z_{i-1}$ and $\left\{z_{i-1}\right\}$ occur at different levels ${ }^{[14}$ NF inherits this phenomenon. The theory would be inconsistent if it did not, but this becomes the source of many unintuitive results in the one-sorted theory.

In TST without the axiom of infinity, we can form Kuratowski pairs, and $\left\langle x_{i}, y_{i}\right\rangle=$ $\{\{x\},\{x, y\}\}$ occurs two levels above $x_{i}$ and $y_{i}$. In TSTI, one can implement homogeneous (Quine) pairs. Ordered pairs $\left\langle x_{i}, y_{i}\right\rangle$ then occur at level $M_{i}$ and the graphs of the corresponding projection functions are sets. As a result, standard mathematical structures, such as finite products and coproducts, have homogeneous implementations. In particular, one can define a reasonable theory of categories.

Lemma 1.4. In TSTI, $\operatorname{Cat}\left(M_{i}\right) \in_{i} \operatorname{Cat}\left(M_{i+1}\right)$ and $\operatorname{Set}\left(M_{i}\right) \in_{i} \operatorname{Cat}\left(M_{i+1}\right)$, generally, where the object of objects for $\operatorname{Set}\left(M_{i}\right)$ is $V_{i}$.

In this sense, one is able to handle the issue of large and small categories in TST. However, a number of classical adjoint equivalences only exist modulo type-shifting. An example that arises, repeatedly in NF, is the failure of cartesian closure. Both function spaces and binary products exist, but the former are one type higher:

$$
\left(x_{i} \Rightarrow y_{i}\right) \in M_{i+1} \wedge x_{i} \times y_{i} \in M_{i}
$$

As a result, the universal natural transformations are undefined. One can, however, conceive of an external form of evaluation. $\sqrt{15}$

$$
e v_{x, y}^{*}:\left(x_{i} \Rightarrow y_{i}\right) \times \iota \text { " } x_{i} \rightarrow \iota " y_{i}, \text { where }\left\langle f,\left\{z_{i-1}\right\}\right\rangle \mapsto\left\{f\left(z_{i-1}\right)\right\}
$$

The categorical structures definable in TST motivate the need to consider external, "type-shifting" functors between levels. The following summary of results anticipates a number of properties in $\mathcal{N}$, the category of NF sets.

[^12]Lemma 1.5. In TSTI, at a given level $M_{i}$, one can form all finite limits and finite coproducts. Coequalizers can, in general, only be formed over diagrams that are "typeraised" images of diagrams in $M_{i-1}$. Similar results for analogues of the powerobject monad and cartesian closure follow from methods such as ev* above.

### 1.2.2 Canonical Embeddings

We consider an operation to be externally definable if it occurs as a valid comprehension instance of TST(I). In other words, if the action corresponds to an external (i.e. inhomogeneous) substitution operation:

$$
\overrightarrow{a_{i}} \mapsto\left\{y_{j} \mid \Phi\left(\vec{a}_{i}, y_{j}\right)\right\}
$$

where $\Phi\left(\vec{z}_{i}, y_{j}\right)$ is a wff of $\operatorname{TST}(\mathrm{I})$. A functor is externally definable if its actions on objects and morphisms are.

A number of embeddings of $M_{i}$ into $M_{j}$, where $i \leq j$, are externally definable:

1. $\iota_{i}: M_{i} \rightarrow M_{i+1}: x_{i} \mapsto\left\{x_{i}\right\}$.
2. $\iota_{i}{ }^{\prime \prime}: M_{i} \rightarrow M_{i+1}: x_{i} \mapsto\left\{\left\{z_{i-1}\right\} \mid z_{i-1} \in_{i-1} x_{i}\right\}$. This anticipates a pattern where one can inject $M_{i}$ into $M_{i+1}$ by adding brackets around elements once, $k$ levels down, where $k \leq i$.
3. $\iota_{i}^{n}: M_{i} \rightarrow M_{i+n}: x_{i} \mapsto\left\{\ldots\left\{x_{i}\right\} \ldots\right\}$, the n-fold singleton of $x_{i}$.
4. $B_{i}: M_{i} \rightarrow M_{i+2}: x_{i} \mapsto\left\{y_{i+1} \mid x_{i} \in_{i} y_{i+1}\right\}$, which we will call the Boffa Embedding.

The most common form of embedding we will use is the $\iota$ operation. Preservation of structure (e.g. functions mapping to functions) under the $\iota$ operation will require its application a certain number of levels down. $\iota$ corresponds to the $T$-functor in $\mathcal{N}$ (see Section 3.1). The following "meta-theorem" is essential to our development of category theory in both TST(I) and NF ${ }^{16}$

[^13]Theorem 1.6. An externally definable type raising operation $M_{i} \rightarrow M_{i+1}$, taken by applying the ८ operation i levels down, preserves all structure in level $M_{i}$.

As with the canonical model of TST, the idea of applying an operation " $i$ levels down" can be formalized in an elementary topos. The unit of the powerobject monad $\{\cdot\}_{C}$ : $C \rightarrow P C$ corresponds to the $\iota$ operation. $\exists$, the left adjoint to the powerobject functor $(\exists \dashv P)$, corresponds to the "jump" operation, $j(\tau)=\lambda x . \tau$ " $x$.

Definition 1.7. Given a function $\tau$, the jump operator $j$ denotes the evaluation of $\tau$ "one level down."

$$
j(\tau)=\lambda x . \tau^{"} x: x \mapsto\left\{\tau^{\prime} y \mid y \in x\right\}
$$

Iteration of the $j$ operator $n$ times is denoted $j^{n}(\tau)$ where, for example, $j^{2}(\tau)^{‘} x=$ $\{\tau " y \mid y \in x\}$.

Lemma 1.8. In the internal language of a topos, $\exists_{\{ \}_{P^{i-n_{C}}}^{n}}^{n}: P^{i} C \rightarrow P^{i+1} C$ applies the singleton operation $n$ levels down, where $n \leq i$.

## Lateral Functions in TST

The jump-operator $j$ is an example of what we have referred to as an externally definable function. Such functions are essential for any reasonable development of category theory in TST (or NF, for that matter). Our rationale for treating this functions as "synthetic" (i.e. in a purely syntactic manner) is understandle, as they exist in some (generic) metatheory but are not sets of our model.

Nevertheless, while our goal at the outset - at least as far as category theory goes - was to assume as little as possible about the metatheory, we should develop some understanding of external functions as semantic objects. Our means of accomplishing this will be to introduce the idea of lateral functions ${ }^{[7]}$

[^14]Definition 1.9. Given a model $\mathcal{M}$ of TST in some metatheory, we say that a function $F$ of the metatheory is $n$-lateral if $n$ is a positive integer of the metatheory and the restriction (where appropriate) of $F$ to each level $M_{i}$ of $\mathcal{M}$ induces a function defined by the action: $\iota^{n \iota} x_{i} \mapsto F^{\iota} x_{i}$, whose graph is a set in $\mathcal{M}$. Similarly, we say a function $G$ is $-n$-lateral if the graph corresponding to $x_{i} \mapsto \iota^{n ‘} G^{6} x_{i}$ is a set in $\mathcal{M}$.

Notice that a definable external function $G$, corresponding to the formula $\psi\left(x_{i}, y_{j}\right)$ :

$$
y_{j}=G^{6} x_{i} \Longleftrightarrow \psi\left(x_{i}, y_{j}\right)
$$

is simply a $(j-i)$-lateral function.

Example 1.10 (Lateral Functions). Standard operations $\iota$ and $\cup$ are $+/-1$-lateral, respectively. The Boffa embedding is 2-lateral.

## Preview: Lateral Functors, Toward TCT

Lateral functions extend to lateral functors in the expected manner: a functor $F$ is lateral if its respective action maps $F_{0}$ (action on objects) and $F_{1}$ (action on morphisms) are lateral functions ${ }^{18}$

Example 1.11 (The $T$-Functor in TST). The $T$-functor is ubiquitous in the category theory of NF. A variant also exists for TST, clearly induced by $\iota: x \mapsto\{x\}$, whose actions on morphisms and objects are defined:

$$
\begin{aligned}
& T_{0} \equiv \iota_{i}{ }^{"}: M_{i} \rightarrow M_{i+1}: x \mapsto \iota " x \quad \text { (Action on Obj } \\
& T_{1}: \operatorname{Mor}\left(M_{i}\right) \rightarrow \operatorname{Mor}\left(M_{i+1}\right):(f: x \rightarrow y) \mapsto(T(f): \iota " x \rightarrow \iota " y: z \mapsto\{f(z)\})
\end{aligned}
$$

(Action on Morphisms)
Clearly this is just a 1-lateral embedding of $C a t_{i}$ into $C a t_{i+1}$.

Example 1.12 (The Powerset Functor). In the canonical model of TST, $M_{i+1}=P\left(M_{i}\right)$.

[^15]But we can also define the powerset operation locally, as a 1-lateral functor:

$$
\begin{aligned}
& \left(P_{i}\right)_{0}: M_{i} \rightarrow M_{i+1}: x \mapsto P^{\prime} x \\
& \left(P_{i}\right)_{1}: \operatorname{Mor}\left(M_{i}\right) \rightarrow \operatorname{Mor}\left(M_{i+1}\right):(f: x \rightarrow y) \mapsto(P f: s \subset x \mapsto\{w \mid \exists z \in s . f(z)=w\})
\end{aligned}
$$

(Action on Morphisms)

Notice, as $T$ and $P$ are functors with equivalent degrees of inhomogeneity (i.e. both 1-lateral), we can define a natural transformation $\{\cdot\}: T \hookrightarrow P$. In fact, this turns out to be the unit of a relative monad, anticipating the relative powerset monad of NF ${ }^{19}$

Example 1.13 (Presheaf Functor). The standard presheaf functor $\widehat{(-)}: C a t \rightarrow C A T$ is maps each (small) category $(C)$ to its category of (contravariant) presheaves $S e t^{\mathcal{C}^{o p}}$. Given some model $\mathcal{M}$ of TST(I), we can define a lateral presheaf functor, using the standard definition of internal presheaves.

Consider a category $\mathbb{C} \in C a t_{i}$ (i.e. the structure morphisms $d_{0}, d_{1}, i, m$ and the collections of objects and morphisms $C_{0}, C_{1}$ are sets of $\left.M_{i}\right){ }^{20}$ The appropriate functors to consider are those from $\mathbb{C}^{o p}$ to $S e t_{i}$. But we know from above that $S e t_{i} \in C a t_{i+1}$, with $\left(\text { Set }_{i}\right)_{0}=M_{i}$, so we cannot speak of (internal) functors from $\mathbb{C}$ to $S e t_{i}$ directly. Instead, we need to consider the internal presheaf category, $\operatorname{Set}_{i}^{\mathbb{C}^{\text {op }}} \cdot{ }^{21}$

Definition. A (contravariant) internal presheaf $F=\left(F_{0}, \gamma_{0}, e\right)$ from $\mathbb{C}^{o p} \rightarrow$ Set $_{i}$ consists of a pair of maps (each of whose graphs is a set in $M_{i}$ ), $\gamma_{0}: F_{0} \rightarrow C_{0}$ and $e: C_{1} \times_{d_{1}} F_{0} \rightarrow F_{0}$ satisfying (contravariant)functoriality conditions:

$$
\begin{array}{lr}
C_{1} \times_{d_{1}} F_{0}=\left\{\langle f, x\rangle \mid d_{1}(f)=\gamma_{0}(x)\right\} & \left(f: c \rightarrow c^{\prime} \text { is a morphism of } \mathbb{C}, x \in F\left(c^{\prime}\right) .\right) \\
\gamma_{0} \circ e=d_{0} \circ \pi_{1} \\
e \circ\left\langle 1, i \circ \gamma_{0}\right\rangle=i d_{F_{0}} & \left(F\left(f: c \rightarrow c^{\prime}\right): F\left(c^{\prime}\right) \rightarrow F(c)\right) \\
e \circ e=e \circ 1 \times m & \left(F\left(1_{c}\right) 1_{F(c)}\right) \\
& (F(g f)=F(f) F(g))
\end{array}
$$

A map $\tau:\left(F_{0}, \gamma_{0}, e\right) \rightarrow\left(G_{0}, \delta_{0}, e^{\prime}\right)$ between interal presheaves is simply a map $\tau: \delta_{0} \rightarrow \gamma_{0}$ in the slice category $S e t_{i} / C_{0}$ (which is itself a category in $C a t_{i+1}$ ) commuting with the action maps $e$ and $e^{\prime}$.

[^16]Thus, the object component of $\widehat{(-)}$ defines a 1-lateral (partial) function from $M_{i}$ to $M_{i+1} \cdot{ }^{[22}$ In other words,

$$
\widehat{(-)_{i}}: C a t_{i} \rightarrow C a t_{i+1}: \mathbb{C} \mapsto S e t_{i}^{\mathbb{C} o p}
$$

Remark. Notice that the role of $\widehat{(-)}$ is analagous to that of $P$ :

$$
\left(P_{i}: M_{i} \rightarrow M_{i+1}: x_{i} \mapsto P^{‘} x_{i}\right) \sim\left({\widehat{(-)_{i}}}_{i}: \text { Cat }_{i} \rightarrow \text { Cat }_{i+1}: \mathbb{C} \mapsto S e t_{i}^{\mathbb{C} p}\right)
$$

Indeed, just as $\iota^{\prime \prime}$ embeds $x_{i}$ into $P^{‘} x_{i}$, the Yoneda Functor embeds $\mathbb{C}$ into $S e t_{i}^{\mathbb{C} o p}$. This is our first hint that the Yoneda Embedding $\mathcal{Y}$ is actually a form of $T$-functor.

Remark (Toward TCT). Given a model $\mathcal{M}$ of TST, our vision of TCT (typed category theory) is something along the lines of an $N$-indexed coproduct, $\coprod_{i \in N} C a t_{i}$, where each category is "small" (i.e. is internal at some level) and each functor is $n$-lateral. A 0 -lateral functor is locally (i.e. when restricted to each level $M_{i}$, where appropriate) internal. Furthermore, any $n$-lateral functor corresponds to a locally internal functor $n$-levels up/down (Definition 1.9).

The temptation is, of course, to derive a single sorted theory of categories from this approach. If $T C T$ is defined as above (where categories must be (locally) internal), then one is looking for a colimiting operation ${ }^{23}$ But the introduction of lateral functions implies that one could take a possibly more direct approach.

A category $\mathcal{C}$ could simply be defined by a class $C_{1}$, where the set of $j$-morphisms corresponds to $C_{1} \cap M_{j}$. But the structure maps would, of course, still be 0-lateral. So any category $\mathcal{C}$ would simply correspond to a coproduct of internal categories $\coprod_{i \in N} \mathbb{C}_{i}$ (where $\mathbb{C}_{j}$ is trivial, whenever $C_{1} \cap M_{j}=\emptyset$ ).

[^17]
### 1.3 An Introduction to NF

New Foundations can be seen as a one-sorted version of TST. One maintains the typing discipline of the syntax, but the sets that are realized by instances of "stratified" comprehension are single sorted. To a category theorist, one might say this feels like a Grothendieck Construction of a single category, fibred over a base $\mathcal{C}$, derived from an indexed category (i.e. from a (pseudo)-functor $\left.\mathcal{C}^{o p} \rightarrow C a t\right)$. Nuance arises, however, from the ability to quantify over all types.

The syntax of NF is not typed, as in TST, but one restricts comprehension to formulae that can be "typed" in such a way as to yield a well formed formula of TST.

Definition 1.14. A formula $\phi(\vec{x})$ in the language of set theory is said to be stratified if there exists a function $t: \operatorname{Var}(\phi) \rightarrow N$ that assigns to each variable ' $x$ ' occurring (free or bound) in $\phi(\vec{x})$, a natural number $t(x)$ such that the following holds for every subformula of $\phi$ :

1. For any occurrence of ' $x=y$ ' in $\phi(\vec{x}), t(x)=t(y)$
2. For any occurrence of ' $x \in z^{\prime}$ ' in $\phi(\vec{x}), t(z)=t(x)+1$

Such a function $t$ is called a stratification of $\phi$.

## Example 1.15.

$$
\begin{aligned}
& \exists x, y, z \cdot x \in y \wedge y \neq z \wedge x \in z \\
& \forall x \exists y \forall w \cdot w \in y \Longleftrightarrow \forall z \cdot z \in w \Longrightarrow z \in x \\
& \exists x \cdot x \in x \\
& \exists x, y, z \cdot x \in y \wedge y \in z \wedge x \in z
\end{aligned}
$$

While NF has a finite axiomatization due to Hailperin (see [16]), the most intuitive description involves Extensionality and the axiom scheme of Stratified Comprehension:

## 1. Extensionality

$$
\forall x, y \cdot x=y \Longleftrightarrow \forall z \cdot z \in x \Longleftrightarrow z \in y
$$

2. Stratified Comprehension For all stratified instances of the scheme:

$$
\forall \vec{x} \exists y \forall z . z \in y \Longleftrightarrow \phi(\vec{x}, z)
$$

where $y$ is not free in $\phi$.

Example 1.16 (Some Sets and Their Stratified Definitions).

$$
\begin{align*}
& \{x, y\}=\forall x, y \exists z \forall w \cdot w \in z \Longleftrightarrow w=x \vee w=y  \tag{Pairing}\\
& P^{‘} x=\forall x \exists y \forall z \cdot z \in y \Longleftrightarrow \forall w \cdot w \in z \Longrightarrow w \in x  \tag{Powerset}\\
& V=\exists y \forall z \cdot z \in y \Longleftrightarrow z=z  \tag{UniversalSet}\\
& B^{\prime} x=\forall x \exists y \forall z \cdot z \in y \Longleftrightarrow x \in z \\
& V-x=\forall x \exists y \forall z . z \in y \Longleftrightarrow z \notin x
\end{align*}
$$

Remark (Unstratified $\neq$ Non-existence). Just as $\exists x . x \in x$ is unstratifed, the Russell class, $\{x \mid x \notin x\}$, does not correspond to a stratified formula, so is not a valid instance of comprehension in NF. But, the fact that $\exists x . x \in x$ is unstratified does not mean that one cannot deduce it as a theorem of NF.

$$
N F \vdash V \in V \Longrightarrow N F \vdash \exists x \cdot x \in x
$$

To make sense of this, we consider a slightly weaker form of stratification.
Definition 1.17. $\phi$ is considered weakly stratified if there is a valid stratification of its bound variables.

In general, for a set $\{x \mid \phi(\vec{z}, x)\}$ to be a set of NF, it is sufficient that $\phi(\vec{z}, x)$ be weakly stratified, where the eigenvariable ' $x$ ' is considered bound.

Example 1.18 (Weakly Stratified Formulae).

$$
\begin{aligned}
& \forall x, y \cdot x \in y \wedge z \in y \wedge z \in x \\
& \forall y, z . y \in z \wedge x \in x
\end{aligned}
$$

Remark (Substitution and Weak Stratification). In an instance of the (stratified) comprehension scheme:

$$
\exists y \forall x \cdot x \in y \Longleftrightarrow \phi(\vec{z}, x)
$$

all variables in $\vec{z}$ are bound, as in the case:

$$
V=\exists y \cdot \forall x \cdot x \in y \Longleftrightarrow x=x
$$

We can think of instances where all variables of $\phi$, other than $x$, are bound as defining "concrete" sets of NF. In practice, however, we tend to encounter instances of the form:

$$
\forall \vec{z} \exists y \forall x \cdot x \in y \Longleftrightarrow \phi(\vec{z}, x)
$$

When we prove something like pairing in NF, in a sense we are constructing an operation $V^{n} \rightarrow V$, where $\vec{x}$ contains $n$ variables. In the case of pairing, we obtain the set $\{x, y\}$ by substituting $x$ and $y$ for $v_{1}$ and $v_{2}$ in $z=v_{1} \vee z=v_{2}$. As $v_{1}$ and $v_{2}$ range over all of $V$, from any concrete set $a$ we obtain the intuitively unstratified set $\{a,\{a\}\}$. Another example, is the weakly stratified set abstract:

$$
\{z \mid z \in x \vee z=x\}
$$

While $z \in x \vee z=x$ requires decorating $x$ with distinct types, $x \cup\{x\}$ does turn out to be a set in NF. What this comes down to is the ability to rename the two occurrences of the free variable $x$, yielding the stratified instance of comprehension:

$$
\forall x, y \exists w \forall z . z \in w \Longleftrightarrow z \in x \vee z=y
$$

As $x$ and $y$ each range over $V$, sets of the form $\{z \mid z \in x \vee z=x\}$ are determined in the special case where the same concrete set is substituted for $x$ and $y .{ }^{24}$

Strictly speaking, this conception of comprehension is purely informal and reduces to the "concrete" form, in all cases. But the idea of substituting concrete sets for free variables in $\phi$ (other than the eigenvariable) has important consequences for NF, as a deductive system.

A similar nuance to set-formation in NF involves closed terms. In a stratification of a formula, multiple occurrences of the same closed term can receive distinct values. For example:

$$
\exists x, y \cdot x \in y \wedge x \in V \wedge y \in V
$$

[^18]is stratified, despite the fact that, if treated as a variable, $V$ would require distinct types. The reason is that sets truly are predicates in extension, in NF. Accordingly, we can rewrite the formula in a way that is clearly stratified:
$$
\exists x, y . x \in y \wedge x=x \wedge y=y
$$

Combining the two notions, one might consider the set $\emptyset \cup\{\emptyset\}$, which can be written out as the set abstract:

$$
\{z \mid z \neq z \vee \forall w \cdot w \in z \Longleftrightarrow w \neq w\}
$$

Thus the pathology is not so deep as one might believe, at first glance.

Nevertheless, care needs to be taken. Consider the closed term (i.e. the set) $N$, the natural numbers, implemented in NF as the Frege naturals. A specific $n$ is also a closed term and an element of $N$. Take $n=1$ :

$$
1 \equiv\{x \mid \exists w \cdot w \in x \wedge \forall y, z \cdot(y \in x \wedge z \in x) \Rightarrow z=y\}
$$

It is no trouble to show the following holds:

$$
\exists x \cdot x \in 1 \wedge \forall z \cdot z \in x \Longleftrightarrow z=1
$$

The set $\{1\}$ clearly exists and is a member of 1 . However, quantifying over generic elements of $N$ is not permitted. Thus, a defining characteristic of the Von Neumann implementation of natural numbers $\forall n \in N .\{m \mid m<n\} \in n L^{25}$ To obtain this property for the Frege implementation, we need to extend NF to include the Axiom of Counting. In the following sections we introduce this axiom in two (equivalent) forms, in the context of "counting" and the existence of infinite strongly cantorian sets.

### 1.3.1 The Natural Numbers

Arguably, the key strength of NF is its admission of large sets. As a result, implementations of mathematical objects are closer to those in naive set theory than their ZF

[^19]counterparts. The natural numbers of NF are the Frege naturals, where ' $n$ ' is the set of all $n$-element sets.

Definition 1.19. Given a set $x$, the successor operation $S$ is defined:

$$
S^{6} x \equiv\{a \cup\{z\} \mid a \in x \wedge z \notin x\}
$$

This definition is stratified and, as the ordered pair $\left\langle x, S^{‘} x\right\rangle$ is homogeneous, the graph of $S$ is a set in NF ${ }^{26}$

Definition 1.20 (The Frege Naturals).

$$
\begin{aligned}
& 0=\{\emptyset\} \\
& 1=S^{\prime} 0=\iota^{\prime \prime} V \\
& \cdots \\
& n+1=S^{\prime} n \\
& \cdots \\
& N=\bigcap\left\{b \mid 0 \in b \wedge \forall x \in b \cdot S^{\prime} x \in b\right\}
\end{aligned}
$$

Hence, we can also define the set of all finite sets: $F i n=\bigcup N$.

The implementation of $N$ is formed by the join (intersection) of a collection of sets that would typically be a class. This is a recurrent theme: set theories with full (stratified) comprehension often take the intersection over a well-ordered structure, where one would expect to use induction (as in the standard $\mathrm{ZF}(\mathrm{C})$ case). The general method, used here to obtain $N$, corresponds to NF's recursion theorem ${ }^{27}$ The Frege naturals satisfy both recursion and the Peano axioms, giving $N$ the universal property of a natural numbers object in $\mathcal{N}$.

[^20]
## Returning to Counting

As mentioned above, a defining property of the (more familiar) Von Neumann naturals is:

$$
\forall n \in N .\{m \mid m<n\} \in n
$$

Clearly, this is also a desirable property for the Frege implementation in NF ${ }^{28}$ The appropriate statement of the above property for the Frege implementation is:

$$
\forall n \in N .\{m \mid 0<m \leq n\} \in n
$$

The first case (where $n=0 \equiv \emptyset$ ) holds vacuously, and we can easily prove what appears to be the necessary induction step:

$$
\forall n \in N .\{m \mid 0<m \leq n\} \in n \Longrightarrow\{m \mid 0<m \leq n+1\} \in n+1
$$

This gives the obvious chain of corollaries:

$$
\begin{array}{ll}
\{m \mid 0<m \leq 1\} \in 1 & (n=1) \\
\{m \mid 0<m \leq 2\} \in 2 & (n=2) \\
\ldots & \\
\{m \mid 0<m \leq k\} \in k & (n=k)
\end{array}
$$

But it does not allow us to prove by induction:

$$
\forall n \in N .\{m \mid 0<m \leq n\} \in n
$$

The reason being that the inference of the above statement from the induction step would require stratification. As $n$ is bound, this clearly fails. Thus, we are in the (somewhat uncomfortable) position of knowing $\{m \mid 0<m \leq n\}$ holds for any concrete $n$ - what Rosser refers to as induction in the "intuitive logic" - but we are unable to provide a formal proof, using the axioms of NF. Furthermore, making this assumption formally (i.e. as a further axiom) is strong ${ }^{29}$

[^21]Definition 1.21 (AxCount I). The Axiom of Counting is the statement: $\forall n \in N .\{m \mid 0<$ $m \leq n\} \in n$.

We give an alternative definition, and prove its equivalence to Definition 1.21, in Section 1.3.4.

### 1.3.2 A Homogeneous Pairing Function

For reasons upon which we will elaborate later, a near necessity for working "practically" in NF is a homogeneous pairing operation. The Quine pairing operation is both homogeneous and surjective, and can be implemented in any (reasonable) set theory with infinity. The premise is simple: if we choose a set $x$, any set either does or does not contain it. By choosing a particularly nice sequence like $\omega=\langle 0,1,2, \ldots\rangle$, we can obtain a unique correspondence between sets and ordered pairs. ${ }^{30}$

Definition 1.22 (Quine's $\theta$ operations).

$$
\begin{aligned}
& \theta_{0}{ }^{‘} x=\{n+1 \mid n \in x \cap N\} \cup\{y \mid y \in x \backslash N\} \\
& \theta_{1}{ }^{‘} x=\theta_{0}{ }^{‘} x \cup\{0\}
\end{aligned}
$$

$\theta_{0}$ and $\theta_{1}$ are homogeneous, so one can form their respective graphs as sets in NF.
Definition 1.23 (The Quine Pair). For two sets $x$ and $y$ :

$$
\langle x, y\rangle=\theta_{0} " x \cup \theta_{1} " y
$$

We obtain the identity: $\left\langle\pi_{1}(z), \pi_{2}(z)\right\rangle=z$ expressed by the commutative diagram:


There is more to be said on the Quine pairing function, but we delay that to Section 5.4. For now, it suffices to observe that we can form sequences $\langle x, y, z\rangle$ as $\langle x,\langle y, z\rangle\rangle$. Thus,

[^22]standard categorical structures such as finite products can be obtained by (homogeneous) finite sequences.

Remark (Implementing Total Functions). Unlike ZF(C), NF can implement total functions as sets. However, when we form the graph of a total function $\{\langle x, f(x)\rangle \mid x=x\}$, what we are really saying is:

$$
\left\{z \mid \pi_{0}(z)=x \wedge \pi_{1}(z)=f(x)\right\}
$$

So we are not free to assign distinct types to ' $x$ ' and ' $f(x)$.' Hence, functions (and relations, more generally) are definable as sets in NF are precisely when their action is stratified (i.e. the syntax describing the action $x \mapsto f(x)$ is stratified) and their graph is homogeneous.

### 1.3.3 Cantor's Theorem

By definition, the universe $V$ is a fixed point for the powerset functor (i.e. $P V=V$ ). In Zermelo-style set theories this would contradict Cantor's theorem, that no set can be in bijection with its powerset. For a claimed bijection $f: x \rightarrow P^{‘} x$, one forms the "paradoxical" subset of $x,\{a \in x \mid a \notin f(a)\}$. The formula defining this set is, however, unstratified. Therefore, one cannot (by stratified comprehension) form the necessary set to contradict Cantor's theorem

What one can prove is a stratified variant of Cantor's theorem, by forming the "paradoxical" set $\{a \in x \mid a \notin f(\{a\})\}$. In other words, NF allows us to prove that there is not a bijection between the sets of singletons of elements of a set $x, \iota " x=\{\{z\} \mid z \in x\}$, and the powerset, $P^{\prime} x$. In fact, $\left|\iota^{\prime \prime} x\right| \lesseqgtr\left|P^{‘} x\right|$. Therefore, we obtain the strange result that $|\iota " V| \leq|V|$. The set of singleton sets is strictly smaller than the universe. Clearly this is not true externally, but any model of NF believes it is the case.

Remark (Rationale for $|\iota " V| \lesseqgtr|V|$ ). This pathology makes more sense, when one observes how it is inherited from TST. In a model of TST, the universe of sets at level

[^23]$T_{i}, V_{i}$, is an object in $T_{i+1}$. If one is to compare cardinalities between a universe and a universe of singletons at a given level, $T_{i+1}$, one must be comparing $\iota$ " $V_{i-1}$ and $V_{i}$. That $\left|\iota " V_{i-1}\right| \leq\left|V_{i}\right|$ is hardly surprising, given the canonical model of TST in Zermelo set theory has $T_{i}=P^{\prime} T_{i-1}$. So, in a sense, the result of NF that $|\iota " V| \lesseqgtr|V|$ can almost be viewed as a direct consequence of the validity of Cantor's theorem (in Zermelo).

### 1.3.4 Cantorian and Strongly Cantorian Sets

In light of the fact that $|\iota " V| \lesseqgtr|V|$, for the "largest" set, it is tempting to think that "small" sets are those which satisfy $|x|=|\iota " x|$. For these sets, Cantor's theorem holds in the classical sense - hence, we refer to them as cantorian. The graph of $\iota$ is not a set in NF and, in fact, it need not exist locally, even for sets which are cantorian. The sets of NF for which the graph of $\iota \upharpoonright a$ is a set and, therefore, a canonical witness to the fact that $a$ is cantorian, satisfy a much stronger condition. We refer to them as strongly cantorian sets.

Definition 1.24. A cantorian set $x$ is one where $|x|=|\iota " x|$. A set $y$ is called strongly cantorian if the graph of $\iota$ :

$$
\iota \equiv\{\langle x,\{x\}\rangle \mid x=x\}
$$

restricted to $y$ (i.e. $\iota \upharpoonright y$ ) is a set.

It is in the folklore of NF that the strongly cantorian sets are the "small" sets. The strongly cantorian sets form a topos subcategory of $\mathcal{N}$ (Theorem 3.64). But, an intuitive understanding of just how "small" these sets are is tricky. On the one hand, given any concrete finite set, we can prove it is strongly cantorian. However, the theorem which states the existence of an infinite strongly cantorian set is strong to NF (i.e. NF + 'there is an infinite strongly cantorian set' implies $\operatorname{Con}(N F)$ ) [9]. It is worth sketching the proof that a concrete finite set is strongly cantorian, as it gives some insight into the nature of set-formation in NF.

Lemma 1.25. Given a concretely finite set $x=\left\{x_{1}, \ldots x_{n}\right\},\{\langle z,\{z\}\rangle \mid z \in x\}$ is a set.

Proof. The following assertion is unstratified:

$$
\forall x \exists y \cdot y=\langle x,\{x\}\rangle
$$

But, the assertion:

$$
\forall x, z \exists y \cdot y=\langle x,\{z\}\rangle
$$

is unproblematic. Therefore, if we substitute a set $w$ for ' $z$ ' and ' $x$ ', we obtain the set $\langle w,\{w\}\rangle$ as a valid set in NF. As $x$ is a concrete set, so are each $x_{i}$, so we can form the ordered pair $\left\langle x_{i},\left\{x_{i}\right\}\right\rangle$ for each element of $x$. Forming a union of finite sets is entirely unproblematic for NF, so we can form the graph of $\iota$ restricted to $x$.

The distinction between quantification over all finite sets and substitution of a given finite set of concrete sets is unintuitive, but has a clear (formal) justification. A more uncomfortable aspect of this result, in the opinion of the author, is its effective imposition of a smallness condition on syntax. A formula in primitive notation can only contain finitely many variables, so one cannot substitute a countably infinite collection of concrete sets to the left of the existential quantifier in the relevant instance of comprehension. The best we can do is form the stratified comprehension instance, for any $n$ :

$$
\forall z_{1}, \ldots, z_{n}, w_{1}, \ldots, w_{n} \exists y \forall t . t \in y \Longleftrightarrow t=\left\langle z_{1}, w_{1}\right\rangle \vee \ldots \vee t=\left\langle z_{n}, w_{n}\right\rangle
$$

By substitution, we can prove any concrete finite set is strongly cantorian. To do any better requires the Axiom of Counting, here stated as an alternative form of Definition 1.21 .

Definition (AxCount II). NF $+\mathbf{A x C o u n t}=\mathrm{NF}+$ 'all finite sets are strongly cantorian.'

The equivalence between $A x C o u n t I$ and $A x C o u n t I I$ requires getting slightly ahead of ourselves and using Definition 1.29: Given a cardinal $\kappa \in N C$ and $x \in \kappa, T \kappa \equiv|\iota " x|$.

For any concrete $n$, we can prove $T n=n$. But quantifying over $N$ formally (i.e. proving the statement: $\forall n \in N . T n=n$ ) requires AxCount. Furthermore, from $\forall n \in N . T n=n$ (all finite sets are cantorian), we can prove a pair of (apparently) stronger statements.

Proposition 1.26 (9)). The following are equivalent:

1. All finite sets are cantorian.
2. All finite sets are strongly cantorian.
3. There is an infinite strongly cantorian set.

Proof. $(2 \Longrightarrow 1)$ is trivial. $(1 \Longrightarrow 3)$ is witnessed by the set $\left\{\left\langle T n, \iota \iota^{\iota} n\right\rangle \mid n \in N\right\}$ and the hypothesis $\forall n \in N . T n=n$. The final step is $(3 \Longrightarrow 2)$ : by induction over $N$, any infinite set has subsets of all finite size - if an infinite set has a subset of size $n$ but not of size $n+1$, it must have $n$ members itself - and it is clear that any subset of a strongly cantorian set is strongly cantorian.

### 1.3.5 Ordinals and Cardinals in NF

We provide a very brief overview of ordinal and cardinal arithmetic in NF, largely summarizing the account given in Forster [9, Section 2.2]. Just as the Frege natural numbers form a set in NF, we can form the sets $N O$ and $N C$ of all ordinals and cardinals, respectively, as sets of equivalence classes.

Definition 1.27. The set of all ordinals $N O$ is defined as the set of equivalence classes of well-orderings under order-isomorphism. The order-type of $N O$ under the canonical ordering relation is denoted by $|N O|_{\leq}=\Omega$.

Definition 1.28. The set of all cardinals $N C$ is the set of $\cong$-equivalence classes.

At least as early as Rosser, it was noticed that $\iota$ would benefit from an analogous action on pairs (hence, relations). This is denoted:

$$
R U S C:\langle x, y\rangle \mapsto\langle\{x\},\{y\}\rangle
$$

Definition 1.29 (T on Ordinals and Cardinals). Given $\kappa \in N C$, we define the cardinal $T \kappa$ as $\mid \iota$ " $x \mid$, for $x \in \kappa$. Likewise, for any $\alpha \in N O, T \alpha$ is the ordinal $\mid R U S C$ " $a \mid$, for
$a \in \alpha, \sqrt{32}$

It is easy to see that $T$ preserves the relevant ordering and cardinality relations. We sometimes refer to this by saying $T$ is monotonic.

Typically, given some cardinal $\kappa$, we define $2^{\kappa}$ as $\left|P^{‘} x\right|$, where $|x|=\kappa$. Such a definition is not homogeneous. Therefore, in NF, it makes more sense to define $2^{(-)}$as a partial function on $N C$.

Definition 1.30. For any cardinal $\gamma \in N C$, the equivalence $\gamma=2^{\kappa}$ is defined if there exists some set $x$ such that $|\iota " x|=\kappa$ and $\left|P^{\prime} x\right|=\gamma$.

Clearly, for $2^{\kappa}$ to be defined, it must be the case that $\kappa \leq T|V|$. In particular, notice $2^{T|V|}=|V|$ and, of course, $2^{|V|}$ is undefined.

It is important to note that, unlike the $T$-operation, $2^{(-)}$is not monotonic. In other words, there exist cardinals $\kappa$ and $\gamma$, such that $\kappa<\gamma$ and $2^{\kappa}=2^{\gamma}$ [18]. In this sense, $2^{(-)}$is weakly monotonic.

Using the homogeneous definition of $2^{(-)}$and the formation of sets by intersection (rather than recursion) allows us to define two operations on cardinals, both of which are of great importance to the study of NF. We will not define the concept in full, but each cardinal gives rise to a Specker Tree [9]. Given a cardinal $\kappa$, we define the set of nodes of the Specker Tree of $\kappa$ as $S p(\kappa)$. This set has a (sort of) converse, which we denote $\Phi(\kappa)$.

Definition 1.31. Given a cardinal $\kappa$, we can define a pair of sets:

$$
\begin{aligned}
& S p(\kappa) \equiv \cap\left\{x \mid \kappa \in x \wedge \forall \gamma \cdot 2^{\gamma} \in x \Rightarrow \gamma \in x\right\} \\
& \Phi(\kappa) \equiv \cap\left\{x \mid \kappa \in x \wedge \forall \gamma \in x .2^{\gamma} \text { exists } \Rightarrow 2^{\gamma} \in x\right\}
\end{aligned}
$$

More informally, we can think of $\Phi(\kappa)$ as the set $\left\{\kappa, 2^{\kappa}, 2^{2^{\kappa}}, \ldots\right\}$.

Notice, in the definition of $\Phi(\kappa), 2^{\gamma}$ is defined only where $\gamma \leq T|V|$. Therefore, we can conceive of cardinals for which application of $\Phi$ results in a finite set. Concrete

[^24]examples are: $\Phi(|V|)=\{V\}$ and $\Phi(T|V|)=\{T|V|,|V|\}$. Moreover, the collection of all such cardinals forms a set, which Forster denotes $S M$.

## Definition 1.32.

$$
S M \equiv\{\kappa \mid \kappa \in N C \wedge \Phi(\kappa) \in F i n\}
$$

We can use induction on $N$ to prove that if $\gamma$ is the $n^{\prime}$ th element of $\Phi(\kappa)$, then $T \gamma$ is the $T n$ 'th element of $\Phi(T \kappa)$ [9]. Thus, we obtain $\Phi(T \kappa) \in$ Fin implies $\Phi(\kappa) \in$ Fin. Equally,

$$
T \kappa \in S M \Rightarrow \kappa \in S M
$$

## $\neg A C$ and Infinity

Arguably, the most important result in the literature of NF is Specker's proof that $N F \vdash$ $\neg A C$ and, by corollary, $N F \vdash \operatorname{Inf}$ [61]. The cardinal arithmetic we have developed above permits a sketch of Specker's argument.

Theorem 1.33 (61]).

$$
N F \vdash \neg A C
$$

Proof. Assume the axiom of choice holds in NF. Using AC, we can choose a minimal cardinal $\kappa$, from the set $S M$, defined above as:

$$
S M \equiv\{\rho \mid \rho \in N C \wedge \Phi(\rho) \in F i n\}
$$

$\Phi(\kappa)$ is a set of $n$ cardinals, the largest of which we refer to as $\gamma . \gamma$ must be such that $2^{(-)}$is undefined, therefore it must be the case that $\gamma \not \leq T|V|$. We obtain the following chain of relations:

$$
\begin{array}{lr}
T|V|<\gamma \leq|V| & (\text { AC and } \gamma \not \leq T|V|) \\
T^{2}|V|<T \gamma \leq T|V| & (T \text { monotone }) \\
T|V|=2^{T^{2}|V|} \leq 2^{T \gamma} \leq 2^{T|V|}=|V| & \left(2^{(-)}\right. \text {weakly monotone) }
\end{array}
$$

As we observed above, $\Phi(T|V|)$ has 2 elements and $\Phi(|V|)$ has 1 . Therefore, in the case that $T|V|=2^{T \gamma}, \Phi(T \gamma)$ has 2 elements; otherwise $\Phi(T \gamma)$ is a singleton set.

We now consider the set $\Phi(T \kappa)$. $T \gamma$ is the $T n^{\prime}$ 'th element of $\Phi(T \kappa)$. As $\gamma$ is the maximal element of $\Phi(\kappa)$, we obtain:

$$
\Phi(T \kappa)=T^{"} \Phi(\kappa) \cup \Phi(T \beta)
$$

Thus the size of $\Phi(T \kappa)$ is either $T n+1$ or $T n+2$.

Notice, we have indirectly obtained the result that $\Phi(T \kappa)$ is finite. As $\kappa$ is the minimal element of $S M$, this implies $\kappa \leq T \kappa$. On the other hand, the earlier result:

$$
\Phi(T \kappa) \in F i n \Rightarrow \Phi(\kappa) \in F i n
$$

gives $\kappa \leq T^{-1} \kappa{ }^{33}$ We conclude $\kappa=T \kappa$ and, therefore, $\Phi(\kappa)=\Phi(T \kappa)$. This implies that either $n=T n+1$ or $n=T n+2$. In either case, we obtain a contradiction ${ }^{34}$

As the axiom of choice holds over finite sets, no model of $N F \vdash \neg A C$ could consist of only finite sets.

Corollary 1.34 ([61]). $N F \vdash \operatorname{Inf}$

For a direct proof that $V$ is an infinite set (also due to Specker), we refer the reader to [9, Theorem 2.2.7]. Finally, we should note that Specker's argument does not hold in NFU, where cardinal arithmetic is not quite the same as NF. In fact, NFU can be extended to include or negate the axiom of choice [19].

## Burali-Forti? The Answer is NO

Classically, a set of all ordinals will lead to the Burali-Forti paradox. Any collection of ordinals is itself well-ordered under the natural ordering relation. The order type of

[^25]the set of all ordinals, $N O$, under the induced natural ordering, is denoted $\Omega$. But, by definition, $\Omega \in N O$. So one might expect to derive a contradiction from the abstract property of orderings: no initial segment of a well-ordered series is order-isomorphic to the entire series ${ }^{35}$ This is the case for Von Neumann ordinals.

The Von Neumann ordinals have the property: given any ordinal $\alpha, \alpha$ is order-isomorphic to $\operatorname{seg}_{<}(\alpha)$, the set of ordinals less than $\alpha$ (under the natural ordering relation). Let $N O$ denote the set of all ordinals. Under the natural ordering, $N O$ has order-type $\Omega$; and $\operatorname{seg}_{<}(\Omega)$, the segment of all ordinals less than $\Omega$, is order-isomorphic to $\Omega$. Thus, in any implementation where $\left|\operatorname{seg}_{<}(\alpha)\right|_{\leq}=|\alpha|_{\leq}$, the well-ordered series of all ordinals is order-isomorphic to an initial segment of itself. This yields the paradox of Burali-Forti.

In NF, however, $\alpha$ and $\operatorname{seg}_{<}(\alpha)$ are inhomogeneous (i.e. cannot receive the same type in any stratification), for any ordinal $\alpha$. Rather, we obtain the result [48]:

$$
\forall \alpha \in N O .\left|\operatorname{seg}_{<}(\alpha)\right|_{\leq}=T^{2} \alpha
$$

In the case of $\Omega$, rather than a paradox, we obtain the harmless result: $T^{2} \Omega \leq \Omega$.

While the proof of Burali-Forti does not go through for $N O$, there are cases where a similar argument precludes the existence of a set, in NF. We employ one such example in Proposition 3.78.

### 1.4 Important Variants of NF

We have framed NF as a single sorted version of TST, but it is just one of a number of set/class theories with an axiom of stratified comprehension. In addition to NF, three particular theories will have relevance for this thesis:

1. KF
2. NFU (+ Infinity) (+ Choice)

[^26]3. ML

### 1.4.1 KF

KF was introduced in [13] and has the following axiomatization:

1. Extensionality
2. Pair-set
3. Powerset
4. Sumset
5. Stratified $\Delta_{0}$-Separation

KF is Mac Lane set theory, with $\Delta_{0}$-Separation restricted to those formulae that are stratified. It is a sub-theory of both NF and Mac Lane set theory.

Lemma 1.35. [13] $K F+\exists y . \forall x . x \in y=N F$

Lemma 1.36. [13] $K F+$ "Every set is strongly cantorian" $=M a c$

In light of the connection between Mac and topos theory, the result that strongly cantorian sets of NF form a topos is unsurprising. ${ }^{36}$ In fact, one might expect the internal language of the category of KF sets to be very similar to that of a topos. But restricting to stratified $\Delta_{0}$-separation turns out to complicate matters more than one might think. Thus, we consider some useful extension of KF, the latter two of which were introduced in the author's collaborative work with Forster and Vidrine. 14] CE, in particular, will turn out to be an important extension for the category theory of both KF and NF.

## Various Extensions of KF:

- KF + Inf: KF + The existence of a Dedekind infinite set

[^27]- $\mathbf{K F}+\mathbf{I O}: \mathrm{KF}+$ "Every set is the same size as a set of singletons"
- KF + CE: KF + "Every family of pairwise disjoint sets is the same size as a set of singletons"

Remark (Implementing Quine Pairs in KF + Inf). One could implement Kuratoski pairs but, practically speaking, working in a stratified theory requires type-level (Quine) ordered pairs. In NFU - another stratified theory (which does not prove infinity) - typelevel pairs exist if and only if there exists an infinite set. In KF, we are only able to prove: Inf $\Longrightarrow$ type-level (Quine) pairing.

Theorem 1.37. $K F+\operatorname{Inf}$ (in the form of a Dedekind-infinite set) has an implementation of type-level ordered pairs.

Proof. Recall that a Dedekind-infinite set $x$ has a proper injection $f: x \mapsto x$. To implement the Quine pairing function, as was described in Section 1.3, one needs to implement $N$. This is achieved by taking any $y \in x$ that is not in the range of $f$. Consider the following intersection $\bigcap\left\{z \subseteq x \mid y \in z \wedge f^{\prime \prime} z \subset z\right\}$. The implementation of $N$ takes $y$ as ' 0 ' and $f$ as the successor operation.

As $x$ is a concrete set, we are able to form homogeneous Quine pairing functions. Furthermore, in the category $\mathcal{K}$, this implementation has the obvious universal property of a natural numbers object.

In Chapter 3, we will show that KFI (KF $+\operatorname{Inf}$ ) allows for the construction of (local) products and function spaces. In standard KF, as we are only able to implement Kuratowski pairs, some products may not exist. More precisely: The set corresponding to a given product $x \times y$ exists, but, due to inhomogeneity, the projection function is not necessarily a set.

### 1.4.2 NFU

NFU is the theory of NF with urelemente ${ }^{37}$ The axiom scheme is virtually identical to NF, but extensionality applies only to sets. Extensionally, all atoms would be (vacuously) equivalent. Thus, it is helpful to introduce a predicate of sethood, set, in order to determine the objects for which equivalence can be determined by exstensionality ${ }^{38}$

Definition 1.38. The axioms of NFU are as follows:

- Empty Set: $\forall x . x \notin \emptyset$
- Definition of Sethood: A set $x$ is either the (unique) emptyset or contains some element.

$$
\forall x \cdot \operatorname{set}(x) \Longleftrightarrow x=\emptyset \vee \exists y . y \in x
$$

Thus, given some model $\mathcal{M}$ of NFU, $y \in \mathcal{M} \wedge \neg \operatorname{set}(y)$ implies $y$ is an atom.

- (Set) Extensionality:

$$
\forall x, y, z . z \in x \Longrightarrow[x=y \Longleftrightarrow \forall w . w \in x \Longleftrightarrow w \in y]
$$

- Stratified Comprehension: Given some stratified formula $\phi$ in which the variable $y$ does not occur free,

$$
\exists y . \forall z . z \in y \Longleftrightarrow \phi
$$

It might appear that the addition of atoms (urelemente) is an insignificant departure from NF. After all, $N F=N F U+\forall y$.set $(y)$. But there are significant differences between the two theories ${ }^{39}$ Unlike NF, NFU has a well developed model theory and is

[^28]known to be consistent. In fact, NFU is no stronger than TST [21].
Also unlike NF, both infinity and choice are independent of NFU. Thus we may consider an important extension of NFU: $N F U+$ Choice + Inf.

As the axiom of choice allows us to subvert typing/homegeneity issues related to quotient sets, the category of NFU sets (assuming Choice and Infinity) obviates the need for axioms $\mathbf{C E}$ and $\mathbf{S C U},{ }^{40}$ From this, the reader would be justified in concluding that were our goal to determine the "best" (stratified) foundation for category theory, $\mathrm{NFU}+$ Choice $+\operatorname{Inf}$ may well be it.

### 1.4.3 ML

ML will serve as the ambient class theory for our development of category theory in NF. While it is a distinct theory on its surface, ML is ultimately little more than the second order language of the meta theory in which a given model of NF exists. Indeed, our reason for using ML is its equiconsistency with NF and the fact that we wish to assume as little as possible about our ambient theory of classes - the point of working in $\mathcal{N}$ is, after all, the possiblity of developing category theory in a "closed" setting (i.e. $C A T \in C A T$ or, specifically, $\operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\mathcal{N}))$.

Nevertheless, we should ensure a formal understanding of the theory, however "implicitly" we might wish to work within it.

As with NFU, ML is a one-sorted theory, but there is a need to distinguish between two types of objects. Where sets are distinguished from atoms (in NFU) by the condition, $x=\emptyset \vee \exists y . y \in x$, sets are distinguished from classes (in ML) by the condition, $\exists y . x \in y$.

Definition 1.39. The axioms of ML are as follows:

- Empty Set: $\forall x . x \notin \emptyset$

[^29]- Extensionality:

$$
\forall x, y \cdot x=y \Longleftrightarrow \forall z \cdot z \in x \Longleftrightarrow z \in y
$$

- Definition of Set-hood: $\forall x \cdot \operatorname{set}(x) \Longleftrightarrow \exists y \cdot x \in y$
- Class Comprehension: For any formula $\phi$ in which $y$ does not occur free,

$$
\exists y \cdot \forall x \cdot \operatorname{set}(x) \Longrightarrow(x \in y \Longleftrightarrow \phi)
$$

As extensionality applies to all classes, we can denote the unique class corresponding to this instance of comprehension as $\{x \in V \mid \phi\}{ }^{[11}$

- Definition of set-variables For each $i \in N$, we asssume countably many setvariables of the form $x^{i}$. In a formula, any superscripted variable (e.g. $x^{i}$ ) can range only over set-objects (i.e. can only be substituted by elements of $V$ ). ${ }^{42}$
- Definition of Stratified* A formula $\psi$ is stratified* if it is stratified in the traditional sense, and each atomic subformula is of the form $x^{i}=y^{i}$ or $x^{i} \in z^{i+1}$.
- Set Comprehension: For any stratified ${ }^{*}$ formula $\psi$, the unique class witnessing comprehension, $\{x \in V \mid \psi\}$, is a set. In other words,

$$
\psi \text { stratified }^{*} \Longrightarrow \operatorname{set}(\{z \in V \mid \psi(z, \vec{x})\})
$$

Remark (Set Comprehension in ML). The axiom of set comprehension has a difficult history. Quine's original version restricted only substitution to sets [46]. In other words, given a stratified formula $\psi(x, \vec{y})$,

$$
\exists z_{1} \cdot y_{1} \in z_{1}, \ldots \exists z_{n} \cdot y_{n} \in z_{n} \vdash \operatorname{set}(\{x \in V \mid \psi\})
$$

Quine's admission of bound variables ranging over all classes is problematic for two reasons. First, it is inconsistent (see [48]). Second, as the set comprehension axiom permits formulae with unrestricted quantification over classes, there is no reason to

[^30]think that sets of ML correspond to sets of NF (i.e. even if consistent, it is not a true "theory of classes" for NF).

The original, correct axiomatization of ML is due to Wang [70]. In addition to Quine's restriction on substitution, Wang's axiom of set comprehension is restricted to formulae $\psi(x, \vec{y})$ that are stratified and where quantification occurs only in the form: $\exists z . x \in z$, $\forall x . \exists z . x \in z \Longrightarrow \psi^{\prime}$, or $\exists x . \exists z . x \in z \wedge \psi^{\prime}$, where $\psi^{\prime}$ denotes a further subformula of $\psi$. It is easy to see that Wang's axiom is equivalent to that given in Definition $1.39{ }^{43}$

## Lateral Functions in ML

In order to develop a semantic understanding of external functors in TST, we introduced lateral functions in Definition 1.9. In order to rigorously define the external functors of NF (i.e. functors $\mathcal{N} \rightarrow \mathcal{N}$ whose graphs are not sets of the model), we need to define lateral functions in ML.

The notion we wish to capture is that, given some class function $F, F$ is $n$-lateral if and only if $\operatorname{set}\left(\left\{\left\langle\iota^{n} x, F^{\iota} x\right\rangle \mid x \in V\right\}\right){ }^{[44}$ The key step is to internalize the iterated $\iota$ operation as a class function in ML. We do so by the standard version of NF/ML recursion (i.e. intersection over "large" sets).

Definition 1.40. The following intersection defines a functional relation $I$, such that $I(n, x)=\iota^{n \iota} x:$

$$
I \equiv \bigcap\{y \mid \forall x \cdot\langle\langle 0, x\rangle, x\rangle \in y \cdot \wedge \cdot\langle\langle n, z\rangle, w\rangle \in y \Longrightarrow\langle\langle n, z\rangle,\{w\}\rangle \in y\}
$$

We write $I_{n}$ to denote the subset of $I$ corresponding to $((\{n\} \times V) \times V) \cap I$.

We use ' ' to denote relational composition,

$$
R \mid S \equiv\left\{\left\langle s_{0}, r_{1}\right\rangle \mid \exists x .\left\langle s_{0}, x\right\rangle \in S \wedge\left\langle x, r_{1}\right\rangle \in R\right\}
$$

[^31]Definition 1.41. A class function $F$ of ML is $n$-lateral (for $n \geq 0$ ) if and only if $\operatorname{set}\left(F \mid I_{n}^{-1}\right)$. Similarly, a class function $G$ is $-n$-lateral if and only if $\operatorname{set}\left(I_{n} \circ G\right)$.

We can use Definition 1.41 to define a $n$-lateral functor $\mathcal{N} \rightarrow \mathcal{N}$, by considering the internal functor category $[\mathcal{N}, \mathcal{N}]$ in the category of ML classes ${ }^{45}$

Example 1.42 ( $T$ is a 1-lateral functor). The $T$-functor, $T: \mathcal{N} \rightarrow \mathcal{N}$ is defined by the following actions:

$$
\begin{aligned}
& T_{0}=j^{2 \iota} \iota: V \rightarrow V: x \mapsto \iota " x \quad \quad \text { (action on objects) } \\
& T_{1}=j^{\iota} R U S C: F u n \rightarrow \text { Fun }:(f: x \rightarrow y) \mapsto(R U S C(f): \iota " x \rightarrow \iota " y)
\end{aligned}
$$

(action on morphisms)
Clearly, $T_{0}$ is simply a 1-lateral function as: $\operatorname{set}(\{\langle\{x\}, \iota " x\rangle \mid x \in V\})$. Likewise, as Fun is just the set of functional relations with tagged codomain, $T_{1}$ is also 1-lateral.

### 1.5 Appendix: Typical Ambiguity, from TST to NF

In this appendix, we provide a brief overview of Specker's principle of typical ambiguity and the resulting equiconsistency proof between $T S T+A m b$ and NF [62, ${ }^{46}$ Specker's proof of equiconsistency between NF and $T S T+A m b$ shows the connection between TST and NF is much more than syntactic similarity ${ }^{47}$ We provide a brief overview ${ }^{[87}$

## Typical Ambiguity

Definition 1.43. Let $\Gamma$ denote a collection of formulae in $\mathcal{L}_{T S T} . T S T+A m b(\Gamma)$ denotes the extension of $T S T$ by adding axioms of $\phi \Longleftrightarrow \phi^{+}$, for each formula $\phi$ in $\Gamma$, where $\phi^{+}$

[^32]is the result of raising the type index of each variable by one. TST with full ambiguity, $T S T+A m b$, is the extension given where $\Gamma$ is the collection of all sentences expressible in $\mathcal{L}_{T S T}$.

Specker compared this to the duality between lines and points in certain theories of projective plane geometry. Thinking of lines and points as types, ambiguity in TST is a direct generalization, where 'type' corresponds to 'level.'

The rules of first order logic are unaffected by type-raising, and any closed term that occurs as a premise in the proof is defined by a well-formed formula $\psi$, so has a corresponding premise $\psi^{+}$.

The converse, however, is not generally true. We can construct a basic model where it fails. Consider a singleton set, $x=\{y\}$, and construct the canonical model of TST:

$$
\mathcal{M}_{x}=\left\{x, P x, P^{2} x, \ldots\right\} \models T S T
$$

Consider the formula:

$$
\phi^{+}=\exists x_{1}, y_{1} \cdot x_{1} \neq 1 y_{1}
$$

As $\mathcal{M}_{1}$ contains both $\{y\}$ and $\emptyset, \mathcal{M}_{x} \vDash \phi^{+}$. On the other hand, $\mathcal{M}_{0}=\{y\}$. Hence, $\mathcal{M}_{x} \models \neg \phi$ and typical ambiguity fails in $\mathcal{M}_{x}$.

Despite ambiguity not being a theorem of TST, it may seem that one might easily form a model of TST in which it holds. Until the recent work of Holmes, however, no such model had been discovered 49

## Duality in Projective Plane Geometry

For the theory TZT, we have an obvious permutation of the language $\mathcal{L}_{T Z T}$, whereby variables at level ' $i$ ' are mapped to those at level ' $i+1$ '. Permuting the language of a theory in a way that preserves logical connectives formalizes the concept of duality. The stronger property of ambiguity is related to the existence of such a permutation.

[^33]"Duality" for the permutation exchanging the predicates point and line, for projective plane geometry was first observed by Gergonne [15]. Specker generalized this idea to an arbitrary two-sorted theory with a binary, symmetric relation $I$ corresponding to incidence between points and lines. The formal system of axioms for the theory $\mathcal{G}$ is dual in the sense that each axiom is the dual of another.

- Axiom 1. Any two points are incident to a common line.
- Axiom 1d. Any two lines meet at a unique point.
- Axiom 2. There are four points, no three of which are incident to a common line.
- Axiom 2d.There are four lines, no three of which meet at a common point.

Given any theorem $S$, deducible from the axioms, the dual of $S$ under the permutation (point, line), denoted $S^{*}$, is also deducible 50

$$
\mathcal{G} \vdash S \Longleftrightarrow \mathcal{G} \vdash S^{*}
$$

This is not to say, however, that in any model $\mathcal{M}$ of $\mathcal{G}$ :

$$
\mathcal{M} \models S \Longleftrightarrow \mathcal{M} \models S^{*}
$$

A model which does satisfy this property possesses what Specker refers to as a correlation. A correlation is an endomorphism $\pi: \mathcal{M} \rightarrow \mathcal{M}$ preserving the truth of formulae ${ }^{51}$

As to the question of whether every dual system of axioms has some dual model, Specker defines a specific counterexample. In fact, he proves a more general statement: any dual system of axioms from which one can deduce a theorem, $S \Longleftrightarrow \neg S^{*}$, has no dual model. This result motivates the general principle of ambiguity.

[^34]Definition 1.44. Given a dual theory $T$ and a collection of formulae in the language of $T$, given by $\Gamma, T+A m b(\Gamma)$ denotes the extension of $T$ by adding axioms of $S \Longleftrightarrow S^{*}$, for each formula $S$ in $\Gamma$. The theory $T+A m b$ is the extension given where $\Gamma$ is the collection of all formulae expressible in the language.

Lemma 1.45 ( 62,63$])$. A complete theory $T$, with a permutation $(-)^{*}$ of $\mathcal{L}(T)$ such that, for all formulae $S$,

$$
T \vdash S \Longleftrightarrow S^{*}
$$

has a model with a corresponding automorphism (i.e. correlation) ${ }^{52}$

Given a model $\mathcal{M}$ of $T Z T$, an automorphism is a correlation if it is an $\in$-automorphism, in other words, if it is an automorphism preserving membership between levels. The axioms of $T Z T$ imply the levels of such a model are isomorphic copies of one another. The existence of a model of $T Z T$ with an $\in$-automorphism $\sigma$ is equivalent to the existence of a model of NF. Starting with a model $\mathcal{M}_{T}$ of $T Z T$, we obtain a model of $\mathcal{M}_{N}$, whose collection of objects is given by $T_{0}$, and whose membership relation $\eta$ is induced by the of $\mathcal{M}_{T}$ :

$$
\mathcal{M}_{N} \models a \eta b \Longleftrightarrow \mathcal{M}_{T} \models a \in \sigma(b)
$$

An immediate corollary is the fundamental result ${ }_{53}^{53}$
Theorem 1.46 ([62, 63]). TST + Amb is equiconsistent with $N F$.
Remark. TST is to NF what Projective Plane Geometry is to the following one-sorted theory, with a symmetric binary predicate $I$ :

- Axiom 1. For any distinct elements $a$ and $b$, there is a unique element $c$ such that $a I c$ and bIc.
- Axiom 2. There are four elements, no three of which are related (by $I$ ) to a common element.

[^35]Just as there is a correspondence between models of NF and TST $+A m b$, there is a one-to-one correspondence between models of the above one-sorted theory and projective planes with "polarity" (automorphic correlation) 62].

## Chapter 2

## Relative Algebra

Adjoint relationships pervade the theory of categories. An adjunction can be thought of as the "right" generalization of a two-sided inverse [33. While they are more general, adjoint relationships maintain the symmetry of isomorphisms. Relative adjunctions present a further degree of generality, by relaxing adjoint symmetry [68].

In the relative case, adjoint symmetry exists modulo a "mediating" functor ${ }^{\text {D }}$ A familiar example is the Yoneda Extension, where $F$ is $\mathcal{Y}_{\mathcal{A}}$-left adjoint to $\mathcal{B}(F-,-)$, denoted $F_{y_{\mathcal{A}}} \dashv \mathcal{B}(F-,-)$.

$$
\mathcal{B}(F A, B) \cong \widehat{\mathcal{A}}(\mathcal{A}(-, A), \mathcal{B}(F-, B))
$$



The Yoneda Extension suggests a measure of duality between relative adjunctions and Kan extensions. ${ }^{2}$ However, while the relative adjoint exists generally, the left Kan Extension $L_{\mathcal{A}}(F)$ exists only under certain conditions (i.e. cocompleteness of $\mathcal{B}$ ). The weaker property of adjointness relative to $\mathcal{Y}_{\mathcal{A}}$ "approximates" the stronger property that entails the existence of certain colimits in $\mathcal{B}$ of objects in the "higher typed" (i.e. larger)

[^36]category $\widehat{\mathcal{A}}$.

Where $\mathcal{Y}$ plays a role embedding small categories into the category of locally small categories, $T$ often plays an analogous "type-raising" role in $\mathcal{N}$.

### 2.1 A Generalization of Adjunction

We briefly review and compare classical and relative adjunctions.

Definition 2.1. Given functors $F: \mathcal{C} \rightarrow \mathcal{D}$ and $G: \mathcal{D} \rightarrow \mathcal{C}$, we say $F$ is left adjoint to $G$ (and $G$ is right adjoint to $F$ ), if there is a natural bijection:

$$
\mathcal{D}(F A, B) \cong \mathcal{C}(A, G B)
$$

Such a relationship is written:

$$
F \dashv G
$$

An adjunction is equipped with a pair of universal natural transformations:

$$
\begin{align*}
& \eta: 1_{\mathcal{C}} \Rightarrow G F  \tag{unit}\\
& \varepsilon: F G \Rightarrow 1_{\mathcal{D}}
\end{align*}
$$

(co-unit)

The unit and co-unit correspond to the image of $i d_{F A}: F A \rightarrow F A$ and $i d_{G B}: G B \rightarrow$ $G B$, respectively, under the natural isomorphism of hom-sets defining the adjunction $3^{3}$ Their components correspond to the universal arrows, in the following diagrams:


(co-unit)

[^37]An equivalent characterization of $F \dashv G$ is given by the triangle equalities, emphasizing adjoint symmetry.



## Relative Adjunctions

A relative adjunction involves three functors and possesses "half" of the symmetry expressed by an adjoint relationship:


Definition 2.2. We say $F{ }_{J} \dashv G, F$ is $J$-left adjoint to $G$, if there is a natural bijection:

$$
\Psi_{A, B}: \mathcal{C}(F A, B) \cong \mathcal{D}(J A, G B)
$$

Definition 2.3. We say $G \dashv_{J} F, F$ is $J$-right adjoint to $G$, if there is a natural bijection:

$$
\Phi_{B, A}: \mathcal{C}(B, F A) \cong \mathcal{D}(G B, J A)
$$

Partial symmetry of $F{ }_{J} \dashv G$ and $G \dashv_{J} F$ is expressed by the relative unit and co-unit.
Definition 2.4. Given a J-left adjoint, $F{ }_{J} \dashv G$, the relative unit is the natural transformation $\zeta: J \Rightarrow G F$, whose components are the universal arrows:


Definition 2.5. Given a J-right adjoint, $G \dashv_{J} F$, the relative co-unit is the natural transformation $\theta: G F \Rightarrow J$, whose components are the universal arrows:


Lemma 2.6. [68] In the case of $G \dashv_{J} F, F$ determines $G$ uniquely, but is only determined uniquely in the case that $J$ is co-dense. Dually, where $F{ }_{J} \dashv G, G$ is determined by $F$ only in the case that $J$ is dense.

One can think of the mediating functor, $J$, as "identifying" the objects of the category $\mathcal{D}$, which possess the universal arrows that would normally define (half of) an adjunction. In either case, if $J$ is restricted to the identity functor $1_{\mathcal{D}}: \mathcal{D} \rightarrow \mathcal{D}$, we recover a standard adjoint relationship. However, this thought does not extend as far as one might be tempted to think. If $J$ is the inclusion of a subcategory $\overline{\mathcal{D}} \hookrightarrow \mathcal{D}$, a $J$-relative adjunction, say $F_{J} \dashv G$, need not restrict to an adjunction between $\mathcal{C}$ and $\overline{\mathcal{D}}$. There is no guarantee that the image of a given object $A \in \mathcal{C}$ under $G$ is in the subcategory $\overline{\mathcal{D}}$. This turns out to be the case for a number of important relative adjunctions in $\mathcal{N}$, where $T: \mathcal{N} \rightarrow \mathcal{N}$ serves as the mediating functor, embedding $\mathcal{N}$ into itself.

Example 2.7 ((Relative) Coequalizers in NF). A category $\mathcal{C}$ has (functorial) coequalizers when there is an adjunction, $G \dashv \Delta$, where $G: \mathcal{C} \rightrightarrows \rightarrow \mathcal{C}$ is left adjoint to the functor $\Delta: \mathcal{C} \rightarrow \mathcal{C} \rightrightarrows^{\cdot}$, sending an object $C \in \mathcal{C}$ to the constant functor, $\Delta_{C}: \cdot \rightrightarrows \cdot \rightarrow \mathcal{C}$. The natural isomorphism:

$$
\mathcal{C}(G(f, g), C) \cong \mathcal{C} \rightrightarrows \cdot\left(\langle f, g\rangle,\left\langle 1_{C}, 1_{C}\right\rangle\right)
$$

corresponds to the universal property of coequalizers.
Definition 2.8. The coequalizing functor coeq : Set ${ }^{\cdot} \rightarrow$ Set is defined by the following action, for any $f, g: X \rightarrow Y$ :

$$
\operatorname{coe} q(f, g)=\left\{[y]_{\sim} \mid y \sim z \leftrightarrow \exists x . f(x)=y \wedge g(x)=z\right\}
$$

The canonical morphism $c(f, g): Y \rightarrow \operatorname{coeq}(f, g)$ is defined by $y \mapsto[y]_{\sim}$.

While coeq remains a functor in NF , the associated coequalizing map $c(f, g)$ is inhomogeneous. As $y$ is one type below $[y]_{\sim}$, we cannot form the unit of coeq $\dashv \Delta$. We can, however, form the map:

$$
c(T f, T g): T Y \rightarrow \operatorname{coeq}(f, g)
$$

satisfying the universal (semantic) property of a coequalizer for $T f$ and $T g . c(T f, T g)$ is a component of the relative unit for coeq $T_{T \rightrightarrows} \dashv \Delta$, where coeq is $T^{\cdot-}$-left adjoint to $\Delta$ and $T^{\rightrightarrows}: \mathcal{N} \rightrightarrows \rightrightarrows \mathcal{N} \rightrightarrows$ is the functor given by post-composition with $T$. We obtain a natural bijection:

$$
\mathcal{N}(\operatorname{coeq}(f, g), C) \cong \mathcal{N}^{\cdot \rightrightarrows \cdot}\left((T f, T g), \Delta_{C}\right)
$$

Thus, $\operatorname{coe} q(f, g)$ gives a syntactic choice among the isomorphism class of coequalizers of $(T f, T g){ }_{4}^{4}$

There is a second relative adjunction (Proposition 3.36) associated with coequalizers in $\mathcal{N}$, which coheres with $\bar{G}_{T^{: 马}} \dashv \Delta$ in the sense that they paste to form a symmetric lift.

### 2.1.1 Composition of Relative Adjoints

The principal example of relative adjointness in $\mathcal{N}$, modified-cartesian closure, requires a more general study of the interaction between relative and standard adjunctions. The following results expand upon our initial work in [14].

Lemma 2.9. 14 For any four functors $J, G, F, H$ with $F{ }_{J} \dashv G$ and $H \dashv F$, composition of adjoints with relative adjoints yields another relative adjunction, $H F_{J} \dashv G F$.

Here, we prove a slightly stronger result, making no assumption of a common $J$-left and right adjoint.

Lemma 2.10. The composition of a relative adjunction $F{ }_{J} \dashv G$ with an adjunction $H \dashv K$ is a relative adjunction $H F_{J} \dashv G K$.


[^38]Proof. We denote the co-unit and unit of $H \dashv K$ as $\varepsilon: H K \Rightarrow 1_{\mathcal{E}}$ and $\eta: 1_{\mathcal{D}} \Rightarrow K H$. The relative unit is denoted $\iota: J \Rightarrow G F$. Each component $\iota_{A}$ is given by $\Psi\left(i d_{F A}\right)$, the adjoint transpose of $i d_{F A}$, where $\Psi$ denotes the natural isomorphism:

$$
\Psi_{A, D}: \mathcal{D}(F A, D) \cong \mathcal{C}(J A, G D)
$$

The goal is to prove the existence of a natural isomorphism:

$$
\hat{\Psi}_{A, E}: \mathcal{E}(H F A, E) \cong \mathcal{C}(J A, G K E)
$$

and the relative unit $\hat{\iota}_{A}=\hat{\Psi}\left(i d_{H F A}\right)$.

Define $\hat{\iota}$ as $G \eta_{F} \circ \iota$. This determines the action of $\hat{\Psi}$ :

$$
\begin{aligned}
& \hat{\Psi}:(f: H F A \rightarrow E) \mapsto\left(G K(f) G \eta_{F A \iota_{A}}: J A \rightarrow G K E\right) \\
& \hat{\Psi}^{-1}:(g: J A \rightarrow G K E) \mapsto\left(\varepsilon_{E} H \Psi^{-1}(g): H F A \rightarrow E\right)
\end{aligned}
$$

By definition, $\hat{\iota}=\hat{\Psi}(i d)$. It remains to confirm that $\hat{\Psi}$ is the appropriate natural isomorphism.

$$
\begin{array}{lr}
\hat{\Psi}^{-1} \hat{\Psi}(f)=\varepsilon_{E} \circ H \Psi^{-1}\left(G K(f) \circ G \eta_{F A} \circ \iota_{A}\right) & \\
=\varepsilon_{E} \circ H \Psi^{-1}\left(\Psi\left(K(f) \circ \eta_{F A}\right)\right) & (\Psi=G(-) \circ \iota) \\
=\varepsilon_{E} \circ H\left(K(f) \circ \eta_{F A}\right) & \\
=f \circ \varepsilon_{H F A} \circ H \eta_{F A} & \text { (naturality) }  \tag{naturality}\\
=f & \text { (triangle identity) }
\end{array}
$$

In the other direction:

$$
\begin{array}{lr}
\hat{\Psi} \hat{\Psi}^{-1}(g)=G K \varepsilon_{E} \circ G K H \Psi^{-1}(g) \circ G \eta_{F A} \circ \iota_{A} & \\
=\Psi\left(K \varepsilon_{E} \circ K H \Psi^{-1}(g) \circ \eta_{F A}\right) & (\Psi=G(-) \circ \iota) \\
=\Psi\left(K \varepsilon_{E} \circ \eta_{K E} \circ \Psi^{-1}(g)\right) & \text { (naturality) } \\
=\Psi\left(\Psi^{-1}(g)\right)=g & \text { (triangle identity) }
\end{array}
$$

Naturality is inherited from the standard and relative adjoints.

The dual form of Lemma 2.10 is stated without proof.

Lemma 2.11. The composition of a relative (right) adjunction $F \dashv_{J} G$ and an adjunction $H \dashv K$ is a relative (right) adjunction $F H \dashv_{J} K G$.


### 2.2 Symmetric Lifts

The motivating distinction between a relative adjoint and the special case we refer to as a symmetric lift is:

- Relative Adjoints generalize (one of) the universal natural transformations defining an adjunction.
- Symmetric Lifts generalize (one of) the triangle identities corresponding to an adjunction.

In the classical case, of course, these are equivalent classifications of an adjoint relationship. In the relative case, however, the latter requires stronger assumptions. Nevertheless, symmetric lifts arise naturally in a number of contexts. In $\mathcal{N}$, almost any relative adjunction gives rise to a "partner" that completes a symmetric lift, in the sense that they form a pair of relative adjunctions that paste along a common diagonal to yield a natural isomorphism.

In a more general context, we can prove that any functor between (locally) small categories gives rise to a canonical symmetric lift, approximating the Yoneda Extension diagram 5 The relationship between Yoneda Extensions and symmetric lifts provides

[^39]further intuition for the latter: a symmetric lift forms a "best approximation" of a (left) adjoint to a particular functor which, when it exists, arises as form of (co)completion ${ }^{6}$

Definition 2.12. A pair of relative adjoints, $F_{0} J_{0} \dashv G$ and $F_{1} \dashv_{J_{1}} G$, form a symmetric $l i f \square^{7}$ if the respective relative unit and co-unit paste along the diagonal to yield a natural isomorphism:


When the pasting result holds strictly (i.e. to identity), we will refer to the structure as an exact symmetric lift.

Definition 2.13. A symmetric lift is said to be left-adjointed if $F_{0}$ and $F_{1}$ admit left adjoints $H_{0} \dashv F_{0}$ and $H_{1} \dashv F_{1}$, respecting the commutativity of the diagram, in the sense:

$$
H_{1} \circ J_{1} \cong J_{0} \circ H_{0}
$$

### 2.2.1 Yoneda Extensions as Symmetric Lifts

The Yoneda Extension, the special case of a (left) Kan Extension along the Yoneda Embedding, is a central construct of category theory. Despite the availability of textbooks referencing this structure, it is worth providing a reasonably thorough account.

[^40]We place particular emphasis on the parallel roles of the $T$-functor in $\mathcal{N}$ and the Yoneda Extension, both of which occur in the context of symmetric lifts. ${ }^{8}$

## The Yoneda Extension in Detail

We will speak of the Yoneda Extension, referring not only to the specific left Kan Extension (i.e. $L_{\mathcal{A}}(F)$ in Definition 2.14), but to the broader diagram, describing the Yoneda Embedding as a free co-completion.

Definition 2.14. What we refer to as a Yoneda Extension is the diagram below, where $\widehat{F}$ denotes the left Kan Extension of $\mathcal{Y}_{\mathcal{B}} \circ F$ along $\mathcal{Y}_{\mathcal{A}}$. $L_{\mathcal{A}}(F)$ denotes the left Kan Extension of $F$ along $\mathcal{Y}_{\mathcal{A}}$, although its existence typically requires co-completeness of $\mathcal{B}$.


Remark (Interpreting Yoneda Extensions as Symmetric Lifts). The content of the statement that the Yoneda Extension forms a symmetric lift is: while $\mathcal{B}(F-,-)$ may or may not have an explicit left adjoint $L_{\mathcal{A}}(F)$, it always has a pair of canonical relative adjoints, each of which corresponds to one half of the symmetry obtained in the presence of an actual left adjoint (i.e. the (relative) unit and co-unit).

Remark (Relevance to the Study of NF). Proving that Yoneda Extensions form symmetric lifts is largely a matter of gluing together known results. The existence of a relative adjunction $F y_{\mathcal{A}} \dashv \mathcal{B}(F-,-)$ and the commutativity of the overall diagram are well-known. All that remains is to prove the commutativity of the square corresponds to the pasting of a second relative adjunction. ${ }^{?}$ But doing so is not entirely straightforward, nor is it solely pedagogical.

One of our goals is to develop an understanding of $\mathcal{N}$ as one instance of a more general structure. In doing so, we often seek to associate $T$-relative structures in $\mathcal{N}$ with ap-

[^41]propriate analogues in "structured" categories, whose objects are themselves categories (e.g. Grp, Pos, Cat, etc.). As such, we need to consider that the structure in $\mathcal{N}$, from which we seek to infer something more general, may itself be a degenerate case on account of the "unstructured" nature of sets. A better way to think of this is to view sets as discrete categories: we are trying to identify a general 2-categorical structure, from a particular case in a category with trivial (i.e. only identity) 2-cells. As such objects lack the internal structure from which we might conceive of one object as the (co)completion of another, exchanging free co-completion (i.e. Yoneda Extensions) for a more general property (i.e. symmetric lifts) may be preferable, in studying the relationship between $T$ and $\mathcal{Y}$.

## The $\mathcal{Y}_{\mathcal{A}}$-Left Adjoint

Given a functor $F: \mathcal{A} \rightarrow \mathcal{B}$, there is a relative $\mathcal{Y}_{\mathcal{A}}$-left adjunction $F \mathcal{y}_{\mathcal{A}} \dashv \mathcal{B}(F-,-)$, where $\mathcal{B}(F-,-): \mathcal{B} \rightarrow \widehat{\mathcal{A}}$ is defined as $B \mapsto \mathcal{B}(F-, B): \mathcal{A}^{o p} \rightarrow$ Set:


The natural isomorphism defining the relative adjunction is simply the content of the Yoneda Lemma:

$$
\widehat{\mathcal{A}}(\mathcal{A}(-, A), \mathcal{B}(F-, B)) \cong \mathcal{B}(F A, B)
$$

The relative unit can be defined pointwise as:

$$
\iota_{A}: \mathcal{A}(-, A) \rightarrow \mathcal{B}(F-, F A) ;\left(g: A^{\prime} \rightarrow A\right) \mapsto\left(F g: F A^{\prime} \rightarrow F A\right)
$$

Remark (The Discrete Case). Consider the corresponding diagram in Set, where F: $\mathcal{A} \rightarrow \mathcal{B}$ is just a map $f: A \rightarrow B$, and $\widehat{\mathcal{A}}$ is the powerset $P A$. In this context, $\mathcal{B}(F-,-):$ $\mathcal{B} \rightarrow \widehat{\mathcal{A}}$ is defined by the action: $b \mapsto\{a \mid f(a)=b\}$ (in other words, the map is just $\left.f^{-1}: B \rightarrow P A\right)$. The point of observing the trivial case is, despite the existence of a canonical map $f^{-1}: B \rightarrow P A$, one does not obtain a canonical map $f^{\prime}: P A \rightarrow B$. While there is an obvious map $P f: P A \rightarrow P B$ (it even satisfies $P f \circ\{\cdot\}_{A}=\{\cdot\}_{B} \circ f$ ), one
cannot reduce to a map $P A \rightarrow B$ without some canonical choice method (i.e. structure) on $P B$.

## The Standard Left Adjoint

In the more general case, we confront a similar issue. As presheaf categories are cocomplete, for any functor $F: \mathcal{A} \rightarrow \mathcal{B}$, we can define $\widehat{F}: \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{B}}^{10}$ as the left Kan Extension of $\mathcal{Y}_{\mathcal{B}} \circ F: \mathcal{A} \rightarrow \widehat{\mathcal{B}}$ along $\mathcal{Y}_{\mathcal{A}}, 11$ However, just as we could not determine a canonical map $f^{\prime}: P A \rightarrow B$, we do not obtain the general existence of a left adjoint to $\mathcal{B}(F-,-)$, without assuming further "structure." If $\mathcal{B}$ is co-complete, the left adjoint exists, and is defined pointwise ${ }^{12}$

Definition 2.15. Given a functor $F: \mathcal{A} \rightarrow \mathcal{B}, L_{\mathcal{A}}(F): \widehat{\mathcal{A}} \rightarrow \mathcal{B}$ is defined pointwise:

$$
L_{\mathcal{A}}(F)(H)=\operatorname{colim}\left(F \circ \operatorname{dom}: \mathcal{Y}_{\mathcal{A}} \downarrow H \rightarrow \mathcal{B}\right)
$$

The natural isomorphism of the adjunction $L_{\mathcal{A}}(F) \dashv \mathcal{B}(F-,-)$, defined for $B$ and $H$, is:

$$
\mathcal{B}\left(L_{\mathcal{A}}(F)(H), B\right) \cong \widehat{\mathcal{A}}(H, \mathcal{B}(F-, B))
$$

Using the Yoneda Lemma, each object $\tau: \mathcal{A}(-, A) \rightarrow H$ of the comma category $\mathcal{Y}_{\mathcal{A}} \downarrow H$ corresponds uniquely to an element $\hat{\tau}=\tau_{A}\left(i d_{A}\right)$ of the set $H(A)$. Given a morphism $g: A \rightarrow A^{\prime}$ in $\mathcal{A}$, a morphism $\mathcal{A}(-, g): \tau \rightarrow \theta$ in the comma category corresponds uniquely to the equality:

$$
H(g)(\hat{\theta})=\hat{\tau}
$$

As $\mathcal{Y}_{\mathcal{A}}$ is full and faithful, the correspondence defines an equivalence of categories:

$$
\mathcal{Y}_{\mathcal{A}} \downarrow H \cong E l t s(H)
$$

The natural isomorphism defining the adjunction in Definition 2.15 is an equivalence between co-cones of $F \circ$ dom over $B$ and natural transformations $H \Rightarrow \mathcal{B}(F-, B)$. Given a

[^42]co-cone $(B, t)$, where $t_{\theta}: F A^{\prime} \rightarrow B$ is the leg corresponding to the natural transformation $\theta: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow H$, we obtain a natural transformation $\bar{t}: H \rightarrow \mathcal{B}(F-, B)$ :

where $\bar{t}_{A^{\prime}}(\hat{\theta})=t_{\theta}: F A^{\prime} \rightarrow B$.

The commutativity condition of the natural transformation is therefore:

$$
t_{\theta} \circ F g=\bar{t}_{A}(H g(\hat{\theta}))=\bar{t}_{A}(\hat{\tau})=t_{\tau}
$$

Hence, there is a unique correspondence between natural transformations $H \rightarrow \mathcal{B}(F-, B)$ and co-cones of $F \circ$ dom over $B$. This is easily seen to form a natural isomorphism.

## The $\mathcal{Y}_{\mathcal{B}}$-Right Adjoint

Proving the existence of a second relative adjunction $\widehat{F} \dashv_{\mathcal{Y}_{\mathcal{B}}} \mathcal{B}(F-,-)$ is a straightforward restriction of the adjoint $\widehat{F} \dashv F^{*}$ to representable $\mathcal{B}$-presheaves.


Lemma 2.16. The above diagram forms a $\mathcal{Y}_{\mathcal{B}}$-right adjunction $\widehat{F} \dashv_{\mathcal{Y}_{\mathcal{B}}} \mathcal{B}(F-,-)$. Given an object $B$ and a presheaf $H$, we obtain the natural isomorphism:

$$
\widehat{\mathcal{B}}(\widehat{F}(H), \mathcal{B}(-, B)) \cong \widehat{\mathcal{A}}(H, \mathcal{B}(F-, B))
$$

Proof. $\widehat{F}$ is the left Kan Extension of $\mathcal{Y}_{\mathcal{B}} \circ F$ along $\mathcal{Y}_{\mathcal{A}}, L_{\mathcal{A}}\left(\mathcal{Y}_{\mathcal{B}} \circ F\right)=L_{\mathcal{A}}(\mathcal{B}(-, F-))$ :

$$
\widehat{F}(H)=\operatorname{colim}\left(\mathcal{B}(-, F-) \circ \operatorname{dom}: \mathcal{Y}_{\mathcal{A}} \downarrow H \rightarrow \mathcal{A} \rightarrow \widehat{\mathcal{B}}\right)
$$

As with the Kan Extension in Definition 2.14, $\widehat{F}$ has a right adjoint:

$$
\widehat{\mathcal{B}}(\mathcal{B}(-, F-),-): \widehat{\mathcal{B}} \rightarrow \widehat{\mathcal{A}}
$$

Given some presheaf $G$, the following equivalence follows from the Yoneda Lemma, applied pointwise:

$$
\widehat{\mathcal{B}}(\mathcal{B}(-, F-), G)=G F=F^{*} G
$$

This yields the natural isomorphism, corresponding to $\widehat{F} \dashv F^{*}$ :

$$
\widehat{\mathcal{B}}(\widehat{F}(H), G) \cong \widehat{\mathcal{A}}\left(H, F^{*} G\right)
$$

In the case that $G$ is a representable presheaf $\mathcal{B}(-, B), F^{*} \mathcal{B}(-, B)=\mathcal{B}(F-, B)$. This yields the isomorphism, natural in $B$ and $H$ :

$$
\widehat{\mathcal{B}}(\widehat{F}(H), \mathcal{B}(-, B)) \cong \widehat{\mathcal{A}}(H, \mathcal{B}(F-, B))
$$

defining $\widehat{F} \dashv_{\mathcal{Y}_{\mathcal{B}}} \mathcal{B}(F-,-)$.

The corresponding relative co-unit, $\varepsilon: \widehat{F} \circ \mathcal{B}(F-,-) \rightarrow \mathcal{Y}_{\mathcal{B}}$ can be defined componentwise. Given any object $B$ of $\mathcal{B}$ :

$$
\varepsilon_{B}: \widehat{F}(\mathcal{B}(F-, B)) \rightarrow \mathcal{B}(-, B)
$$

arises as the unique factorization through the colimit $\widehat{F}(\mathcal{B}(F-, B))$ of a canonical cocone of $\mathcal{B}(-, F-)$ odom over $\mathcal{B}(-, B)$, which we denote $(\mathcal{B}(-, B), \gamma):^{13}$


Each leg of the co-cone is a natural transformation between representable presheaves, corresponding to an object of the comma category $\mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{B}(F-, B)$. Thus, given a natural transformation $\tau: \mathcal{A}(-, A) \rightarrow \mathcal{B}(F-, B)$, we define the corresponding leg $\gamma_{\tau}$, of the co-cone $(\mathcal{B}(-, B), \gamma)$ by the identification:

$$
\hat{\gamma}_{\tau}=\hat{\tau}=\tau_{A}\left(i d_{A}\right): F A \rightarrow B
$$

[^43]Thus, given some map $h \in \mathcal{B}\left(B^{\prime}, F A\right)$ :

$$
\gamma_{\tau}(h)=\hat{\tau} \circ h
$$

To prove that the diagram $(\mathcal{B}(-, B), \gamma)$ is a co-cone, we invoke the Yoneda Lemma again. In the diagram described above, $\gamma_{\theta} \circ F g \circ-=\gamma_{\tau}$ is equivalent to $\hat{\gamma_{\theta}} \circ F g=\hat{\gamma_{\tau}}$. By the definition of $(\mathcal{B}(-, B), \gamma)$, this condition is satisfied, as $g \circ-: \tau \rightarrow \theta$ is a morphism in $\mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{B}(F-, B)$.

## The Symmetric Lift

Proposition 2.17. The Yoneda Extension, as defined above, forms a symmetric lift. In other words, pasting the relative unit $\iota: \mathcal{Y}_{\mathcal{A}} \rightarrow \mathcal{B}(F-, F-)$ and the relative co-unit $\varepsilon: \widehat{F}(\mathcal{B}(F-,-)) \rightarrow \mathcal{Y}_{\mathcal{B}}$ yields the identity natural transformation $\widehat{F} \circ \mathcal{Y}_{\mathcal{A}} \rightarrow \mathcal{Y}_{\mathcal{B}} \circ F$.

$$
\varepsilon_{F} \circ \widehat{F}(\iota)=1_{\widehat{F} \circ \mathcal{Y}_{\mathcal{A}}}
$$

Proof. Recall, the relative unit $\iota: \mathcal{Y}_{\mathcal{A}} \rightarrow \mathcal{B}(F-, F-)$ is defined by the action of its components:

$$
\iota_{A, A^{\prime}}:\left(f: A^{\prime} \rightarrow A\right) \mapsto\left(F f: F A^{\prime} \rightarrow F A\right)
$$

Given some object $A$ in $\mathcal{A}$, the natural transformation:

$$
\widehat{F}\left(\iota_{A}\right): \widehat{F}(\mathcal{A}(-, A)) \rightarrow \widehat{F}(\mathcal{B}(F-, F A))
$$

corresponds to the unique factorization through the colimit $\widehat{F}(\mathcal{A}(-, A))$, induced by a sub-diagram $\left(\Phi_{F}, \widehat{F}(\mathcal{B}(F-, F A))\right)$ of the colimiting cone $(\Phi, \widehat{F}(\mathcal{B}(F-, F A))$ ), which itself forms a co-cone of the diagram:

$$
\mathcal{B}(-, F-) \circ \operatorname{dom}: \mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{A}(-, A) \rightarrow \mathcal{A} \rightarrow \widehat{\mathcal{B}}
$$

To define $\left(\Phi_{F}, \widehat{F}(\mathcal{B}(F-, F A))\right)$, we first determine the canonical object of $\widehat{\mathcal{B}}$, which is the base of the colimit, $\widehat{F}(\mathcal{A}(-, A))$. The comma category $\mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{A}(-, A)$ has a terminal object, given by the identity transformation:

$$
\mathcal{A}\left(-, i d_{A}\right): \mathcal{A}(-, A) \rightarrow \mathcal{A}(-, A)
$$

Therefore, $\widehat{F}(\mathcal{A}(-, A))$ is just $\mathcal{B}(-, F A)$, the image of the terminal object: $\mathcal{B}(-, F-) \circ$ $\operatorname{dom}\left(\mathcal{A}\left(-, i d_{A}\right)\right)$. Indeed, given $\tau: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow \mathcal{A}(-, A)$, corresponding by Yoneda to some $\hat{\tau}: A^{\prime} \rightarrow A$, the component $\Psi_{\tau}$ of the colimiting cone $(\Psi, \mathcal{B}(-, F A))$ at $\tau$ is defined:

$$
\Psi_{\tau}=\mathcal{B}(-, F \hat{\tau}) \sim \hat{\Psi}_{\tau}=F \hat{\tau}
$$

$\widehat{F}(\mathcal{B}(F-, F A))$, meanwhile, is the base of a colimiting cone $(\Phi, \widehat{F}(\mathcal{B}(F-, F A)))$ of:

$$
\mathcal{B}(-, F-) \circ \operatorname{dom}: \mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{B}(F-, F A) \rightarrow \mathcal{A} \rightarrow \widehat{\mathcal{B}}
$$

Any natural transformation $\tau: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow \mathcal{A}(-, A)$ determines a natural transformation $\tau_{F}: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow \mathcal{B}(F-, F A)$, corresponding to $\hat{\tau}_{F}=F(\hat{\tau})$. Given some $g: A \rightarrow A^{\prime}$ :

$$
\begin{gathered}
\mathcal{A}\left(A^{\prime}, A^{\prime}\right) \xrightarrow{\tau_{F, A^{\prime}}} \mathcal{B}\left(F A^{\prime}, F A\right) \\
-\circ g \mid \\
\downarrow \\
\mathcal{A}\left(A, A^{\prime}\right) \xrightarrow[\tau_{F, A}]{ } \mathcal{B}(F A, F A) \\
\tau_{F, A}(g)=\hat{\tau}_{F} \circ F g=F(\hat{\tau} \circ g)=F\left(\tau_{A}(g)\right)
\end{gathered}
$$

The sub-diagram $\left(\Phi_{F}, \widehat{F}(\mathcal{B}(F-, F A))\right)$, defined by restricting the colimiting cone $(\Phi, \widehat{F}(\mathcal{B}(F-, F A))$ to edges of the form $\Phi_{\tau_{F}}$, is a co-cone of:

$$
\mathcal{B}(-, F-) \circ \operatorname{dom}: \mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{A}(-, A) \rightarrow \widehat{\mathcal{B}}
$$

The unique factorization of $\left(\Phi_{F}, \widehat{F}(\mathcal{B}(F-, F A))\right)$ through $\widehat{F}(\mathcal{A}(-, A))$ is $\widehat{F}\left(\iota_{A}\right)$.


The second map, $\varepsilon_{F}$, is the co-unit $\varepsilon$, defined as above, restricted to objects in the image of $F$. The component $\varepsilon_{F A}$ arises as the unique factorization, through $\widehat{F}(\mathcal{B}(F-, F A))$, of the co-cone $(\gamma, \mathcal{B}(-, F A))$. As before, given a natural transformation $\alpha: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow$ $\mathcal{B}(F-, F A), \gamma_{\alpha}: \mathcal{B}(-, F A) \rightarrow \mathcal{B}(-, F A)$ is determined by the equivalence:

$$
\hat{\gamma}_{\alpha}=\hat{\alpha}
$$

Given any object of $\mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{A}(-, A), \theta: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow \mathcal{A}(-, A)$ :

$$
\hat{\gamma}_{\theta_{F}}=\hat{\theta}_{F}=F(\hat{\theta})
$$

But this is equivalent to the edge, $\Psi_{\theta}$, of the colimiting cone $(\Psi, \mathcal{B}(-, F A))$. Indeed, the composite:

$$
\varepsilon_{F A} \circ \widehat{F}\left(\iota_{A}\right): \widehat{F}(\mathcal{A}(-, A))=\mathcal{B}(-, F A) \rightarrow \widehat{F}(\mathcal{B}(F-, F A)) \rightarrow \mathcal{B}(-, F A)
$$

defines a factorization:

$$
\varepsilon_{F A} \circ \widehat{F}\left(\iota_{A}\right) \circ \Psi_{\theta}=\gamma_{\theta_{F}}
$$

As factorizations are unique (the universal property of colimits):

$$
\varepsilon_{F} \circ \widehat{F}(\iota)=i d_{\widehat{F} \circ \mathcal{Y}_{\mathcal{A}}}=i d_{\mathcal{Y}_{\mathcal{B}} \circ F}
$$

### 2.3 Relative (Co)Monads

Any adjunction suffices for the construction of a canonical monad and comonad. We recall the basic definitions.

Definition 2.18. A monad $(M, \eta, \mu)$ is comprised of a functor $M: \mathcal{C} \rightarrow \mathcal{C}$ and a pair of natural transformations: ${ }^{14}$

$$
\begin{align*}
& \eta: 1 \Rightarrow M  \tag{unit}\\
& \mu: M^{2} \Rightarrow M
\end{align*}
$$

(multiplication)
satisfying the conditions:

$$
\begin{array}{lr}
\mu \cdot \eta_{M}=\mu \cdot M \eta=1_{M} & \text { (unit law) } \\
\mu \cdot \mu_{M}=\mu \cdot M \mu & \text { (multiplication law) }
\end{array}
$$

[^44]Dually, a comonad $(S, \varepsilon, \delta)$ consists of $S: \mathcal{C} \rightarrow \mathcal{C}$ and a pair of natural transformations:

$$
\begin{align*}
& \varepsilon: S \Rightarrow 1  \tag{co-unit}\\
& \delta: S \Rightarrow S^{2}
\end{align*}
$$

(comultiplication)
satisfying the dual form of the monad laws.

Any (co)monad gives rise to a category of (co)algebras.
Definition 2.19. Given a monad $(M, \eta, \mu)$ on a category $\mathcal{C}, \mathcal{C}^{M}$ denotes the category of $M$-algebras. A $M$-algebra consists of a morphism in $\mathcal{C}, h: M C \rightarrow C$, satisfying the commutative diagrams:


A morphism $f:(X, h) \rightarrow(Y, k)$ between $M$-algebras is a morphism $f: X \rightarrow Y$ in $\mathcal{C}$, such that $f \circ h=k \circ M f$.

Any object in $\mathcal{C}$ gives rise to a free $T$-algebra, $\left(M C, \mu_{C}\right)$.
Example 2.20. Any adjunction $F \dashv G$ yields a canonical algebraic and co-algebraic structure:

$$
\begin{array}{lr}
(G F, \eta, G \varepsilon F) & ((F \dashv G) \text {-induced monad }) \\
(F G, \varepsilon, F \eta G) & ((F \dashv G) \text {-induced comonad })
\end{array}
$$

As relative adjunctions have incomplete symmetry, the most we can state generally is:
Lemma 2.21. [1] A J-left adjunction $F{ }_{J} \dashv G$ gives rise to a relative monad. A J-right adjunction $F \dashv_{J} G$ gives rise to a relative comonad.

As relative monads are developed extensively in [1], we define their dual.
Definition 2.22. A comonad $S: \mathcal{E} \rightarrow \mathcal{C}$ relative to $J: \mathcal{E} \rightarrow \mathcal{C}$ is a triple $(S, \varepsilon,(\hat{)})$, with functor $S$, natural transformation $\varepsilon: S \Rightarrow J$ and a function between the objects of comma categories () : $|S \downarrow J| \rightarrow|S \downarrow S|$. In addition, $(S, \varepsilon, \hat{())}$ satisfies:

1. $\hat{\varepsilon}_{X}=i d_{S X}$
2. For any $X, Y \in|\mathcal{E}|$ and $k: S X \rightarrow J Y, \varepsilon_{Y} \circ \hat{k}=k$
3. For any $k: S X \rightarrow J Y, h: S Y \rightarrow J Z, \hat{h} \circ \hat{k}=\widehat{(h \circ \hat{k})}$

Relative (co)monads directly generalize the definition of (co)monads as (co)extension systems, due to Manes [35]. In the classical setting, a comonad ( $S, \varepsilon, \delta$ ) gives rise to a coextension system $(S, \varepsilon, \hat{( })$, satisfying Definition 2.22 with $J=1_{\mathcal{C}}$ and $\left.\hat{( }\right)=S(-) \circ \delta$. The conditions defining a comonad in "standard" form are equivalent to those defining a coextension system:

$$
\begin{aligned}
& \varepsilon_{Y} \circ \hat{k}=k \sim \varepsilon_{Y} \circ S(k) \circ \delta=k \\
& \hat{h} \circ \hat{k}=\widehat{h \circ \hat{k}} \sim S(h) \circ \delta_{Y} \circ S(k) \circ \delta_{X}=S\left(h \circ S(k) \circ \delta_{X}\right) \circ \delta_{X}
\end{aligned} \quad \text { (multiplication) }
$$

Definition 2.23. An $S$-relative co-algebra consists of an object $X \in|\mathcal{C}|$ and a map $\tau:|X \downarrow J| \rightarrow|X \downarrow S|$ satisfying:

1. $\varepsilon_{Y} \circ \tau(f)=f$
2. $\tau(g \circ \tau(f))=\hat{g} \circ \tau(f)$

A map of relative co-algebras $h:(X, \tau) \rightarrow(Y, \xi)$ is a map $h: X \rightarrow Y$ such that, given any $f: Y \rightarrow J Z, \tau(f \circ h)=\xi(f) \circ h$.

The free relative co-algebras are those of the form $(S X, \lambda k . \hat{k})$ :

$$
\lambda k . \hat{k}:|S X \downarrow J| \rightarrow|S X \downarrow S|
$$

As with comonads, the relative and classical co-algebra conditions coincide when $J=1_{\mathcal{C}}$. In one direction, a co-algebra $j: X \rightarrow S X$ gives rise to a relative co-algebra ( $X, S(-) \circ j)$. In the other, a relative co-algebra $(X, \tau)$ corresponds to the co-algebra $\tau\left(i d_{X}\right): X \rightarrow S X$. The construction of a relative comonad from a $J$-right adjoint, $F \dashv_{J} G$ :

$$
\Psi: \mathcal{D}(F X, J Y) \cong \mathcal{C}(X, G Y)
$$

proceeds as expected.

1. $S=F G$
2. $\varepsilon: F G \rightarrow J$ is given by the relative co-unit
3. $\hat{( })=F(\Psi(-))$

Example 2.24 (The Relative Powerset Monad of NF). Classically, the powerset operation forms a monad in Set. However, the unit and multiplication of this monad, $\{\cdot\}: 1 \rightarrow P$ and $\cup: P^{2} \rightarrow P$, are not homogeneous. Therefore, the analogue in $\mathcal{N}$ is a powerset relative monad. $\sqrt{15}^{15}$

Proposition 2.25. The powerset functor $P: \mathcal{N} \rightarrow \mathcal{N}$ forms a relative monad along $T: \mathcal{N} \rightarrow \mathcal{N}$. The unit $\eta: T \rightarrow P$ is the inclusion $\{\cdot\}: T \rightarrow P$. The object map:

$$
(-)^{\#}:|T \downarrow P| \rightarrow|P \downarrow P| ;(k: T X \rightarrow P Y) \mapsto\left(k^{\#}: P X \rightarrow P Y\right)
$$

is defined by the action:

$$
k^{\#}: S \mapsto\{y \mid \exists x \in S . y \in k(\{x\})\}
$$

Proof. Proving that $P$ is $T$-relative monad requires verifying three conditions.

1. $\{\cdot\}_{X}^{\#}=1_{P X}$. Given any $X \in \mathcal{N},\{\cdot\}_{X}: T X \rightarrow P X$ is just the standard inclusion map. Consider the corresponding action of $\{\cdot\}_{X}^{\#}: P X \rightarrow P X$

$$
\{\cdot\}_{X}^{\#}: S \subset X \mapsto\left\{x \mid \exists x^{\prime} \in S . x \in\left\{x^{\prime}\right\}\right\}
$$

Clearly this is just the identity map on $P X$.
2. $k^{\#} \circ\{\cdot\}_{X}=k$. Given any map $k: T X \rightarrow P Y$, consider the composite action of $k^{\#} \circ\{\cdot\}_{X}:$

$$
k^{\#} \circ\{\cdot\}_{X}:\{x\} \mapsto\{x\} \mapsto\left\{y \mid \exists x^{\prime} \in\{x\} \cdot y \in k\left(\left\{x^{\prime}\right\}\right)\right\}=\{y \mid y \in k(\{x\})\}=k(\{x\})
$$

[^45]3. $h^{\#} \circ k^{\#}=\left(h^{\#} \circ k\right)^{\#}$. Given maps $k: T X \rightarrow P Y$ and $h: T Y \rightarrow P Z$, consider the following actions:
\[

$$
\begin{aligned}
& h^{\#} \circ k:\{x\} \mapsto k(\{x\}) \mapsto\{z \mid \exists y \in k(\{x\}) . z \in h(\{y\})\} \\
& \left(h^{\#} \circ k\right)^{\#}: S \subset X \mapsto\{z \mid \exists x \in S \cdot \exists y \in k(\{x\}) \cdot z \in h(\{y\})\} \\
& h^{\#} \circ k^{\#}: S \subset X \mapsto\{y \mid \exists x \in S \cdot y \in k(\{x\})\} \mapsto\{z \mid \exists x \in S \cdot \exists y \in k(\{x\}) \cdot z \in h(\{y\})\}
\end{aligned}
$$
\]

Thus it is self-evident that $h^{\#} \circ k^{\#}=\left(h^{\#} \circ k\right)^{\#}$.

Among the free (relative) $P$-algebras in $\mathcal{N}$, the obvious one to consider is the fixed point of $P$, the universe object $V$. (-) ${ }^{\#}:|T \downarrow P| \rightarrow|P \downarrow P|$ is the embedding of $T V \Rightarrow V$ into $V \Rightarrow V$. The image of $(-)^{\#}$ defines a bijection between $T V \Rightarrow V$ and the set of functions:

$$
\left\{f: V \rightarrow V \mid \exists g: V \rightarrow V . f=\lambda x . g^{"} x\right\}
$$

This bijection classifies the distributive functions of $\mathrm{NF}{ }^{16}$
Remark (Powerset Monad: NF vs. $\mathrm{ZF}(\mathrm{C})$ ). On the face of it, the asymmetric property (i.e. being a relative monad) satisfied by $P$ in $\mathcal{N}$ is a weakness compared to the standard powerset monad in Set. But what appears to be a weakness can also be seen as a strength.

In some sense, a powerset is a "degenerate" (or, more commonly, "discrete") form of presheaf. Just as objects of Cat are enriched in Set, sets themselves are enriched in the two-element set, $\{\top, \perp\}$. We can extend this idea further:

$$
\begin{array}{lr}
P: S e t \rightarrow \text { Set } \sim(\hat{-}): \text { Cat } \rightarrow C A T & \text { (powersets } \sim \text { presheaf categories) } \\
\{\cdot\}_{A}: A \rightarrow P A \sim \mathcal{Y}_{\mathcal{A}}: \mathcal{A} \rightarrow \widehat{\mathcal{A}} & (\{\cdot\} \sim \text { yoneda embedding) }
\end{array}
$$

Indeed, both $(\widehat{-})$ and $P$ arise as a form of free cocompletion $\sqrt{17}$ But, working over $C a t$, $(\hat{})$ can only be defined as a relative structure. Thus, despite their apparent similarity, $P$ (in Set) and ( $人$ ) (in Cat) correspond to distinct categorical structures.

[^46]In NF, $P$ and $(\widehat{-})$ are on more equal categorical (and foundational) footing . ( $\widehat{-}$ ) does not encounter size restrictions and $P$ does not evade type discipline - both are relative structures. In Chapter 4, we show this is also true of Fam (free coproduct completion), which forms a (internal) relative KZ-pseudomonad in $\mathcal{N}{ }^{18}$

### 2.3.1 A Canonical (Relative) Monad

It is natural to ask: is there a canonical monad one can associate to a given functor $G: \mathcal{D} \rightarrow \mathcal{C}$ ? It is not true, in general. But when $G$ admits a right Kan extension along itself, one obtains the codensity monad of $G, R_{G}(G): \mathcal{C} \rightarrow \mathcal{C}$ (33].

Definition 2.26. A functor $G: \mathcal{D} \rightarrow \mathcal{C}$ is said to be tractable at $C$ when

$$
G \circ \operatorname{cod}: C \downarrow G \rightarrow \mathcal{C}:(a: C \rightarrow U D) \mapsto U D
$$

has a limit $T C$, with limiting cone $\Psi$. If $G$ is tractable at each $C$ in $\mathcal{C}$, it is said to be tractable.

A (pointwise) right Kan extension, $R_{G}(G)$, can also be defined by taking pointwise limits of $G \circ \operatorname{cod}$ :

$$
R_{G}(G)(C)=\underset{\rightleftarrows}{\lim }(G \circ \operatorname{cod}: C \downarrow G \rightarrow \mathcal{D} \rightarrow \mathcal{C})
$$



In other words:

$$
G \text { is tractable } \Longleftrightarrow G \text { admits a right Kan Extension } R_{G}(G)
$$

Proposition 2.27. [35] If $G: \mathcal{D} \rightarrow \mathcal{C}$ is tractable, $R_{G}(G)=T: \mathcal{C} \rightarrow \mathcal{C}$ induces a monad $\left\langle T, \eta,(-)^{\#}\right\rangle$. Furthermore, we obtain a semantic comparison functor between $\mathcal{D}$ and the category of $T$-algebras, $\mathcal{C}^{\mathcal{T}}$.

[^47]Proof. (Sketched.) Given an object $C$ and a morphism $h:(f, D) \rightarrow(g, E)$ in the cocomma category $C \downarrow G$, the limiting cone $\left(T C, \Psi_{C}\right)$ satisfies the commutative diagram:

$C$ itself forms the apex of a cone over $G \circ \operatorname{cod}$, with the objects of the co-slice category as the legs of the cone. Therefore, there is a unique factorization $\eta_{C}: C \rightarrow T C$ through the limiting cone:


Given a morphism $a: X \rightarrow T C$, for any $f: C \rightarrow G D$, define $\Psi_{a}(f)$ as:

$$
\Psi_{a}(f)=\Psi_{X}\left(\Psi_{C}(f) \circ a\right)
$$

where $\left(T X, \Psi_{X}\right)$ is the limit cone over $G \circ \operatorname{cod}: X \downarrow G \rightarrow \mathcal{C}$.
$\left(T X, \Psi_{a}\right)$ is a cone over $G \circ \operatorname{cod}: C \downarrow G \rightarrow \mathcal{C}$, as commutativity is preserved under precomposition. Unique factorization through the limit defines the action of $\left(\begin{array}{l}\# \\ )\end{array}\right.$ on $a$, $a^{\#}: T X \rightarrow T C$. It remains to confirm the monad laws are satisfied by $\eta$ and $(-)^{\#}$.

The semantic comparison functor takes an element $D$ to $\left(G D, \chi_{D}\right)$, where the action of $\chi_{D}$ on a morphism $h: X \rightarrow G D$ is given by:

$$
\chi_{D}:(h: X \rightarrow G D) \mapsto\left(\Psi_{X}(h): T X \rightarrow G D\right)
$$

where $\Psi_{X}(h)$ is the component of the limiting cone $\left(\Psi_{X}, X\right)$ of $G \circ \operatorname{cod}: X \downarrow G \rightarrow \mathcal{C}$ corresponding to $h \in|X \downarrow G|$.

We extend this result and obtain a canonical relative monad $T$ along $J: \mathcal{X} \rightarrow \mathcal{C}$, given any $G: \mathcal{D} \rightarrow \mathcal{C}$ that satisfies the condition of being $J$-tractable.

Definition 2.28. A functor $G: \mathcal{D} \rightarrow \mathcal{C}$ is said to be $J$-tractable along a functor $J: \mathcal{X} \rightarrow \mathcal{C}$, if the functor $G \circ \operatorname{cod}: J X \downarrow G \rightarrow \mathcal{C}$ has a limit for each $X$ in $\mathcal{X}$.

## Proposition 2.29.

1. A J-tractable functor $G: \mathcal{D} \rightarrow \mathcal{C}$, as defined above, forms a canonical J-relative $\operatorname{monad}\left\langle T, \eta,(-)^{\#}\right\rangle$, where $T=\underset{\varliminf}{\lim }(G \circ \operatorname{cod})$.
2. The semantic comparison functor $\Phi: \mathcal{D} \rightarrow J^{T}$ is defined as $\Phi(D)=(G D, \chi)$ :

$$
(f: J X \rightarrow G D) \stackrel{\chi}{\longmapsto}\left(\Psi_{X}(f): T X \rightarrow G D\right)
$$

where $\left(T X, \Psi_{X}\right)$ is the limit cone of $G \circ \operatorname{cod}: J X \downarrow G \rightarrow \mathcal{C}$.
3. Furthermore, if there exists a relative J-left adjoint $F{ }_{J} \dashv G, G F \cong T$ and

$$
\Phi \circ F(X)=\left(T X, \lambda k . k^{\#}\right)
$$

Proof. (1) This is a straightforward generalization of Manes's proof, whereby one can say that $G$ is tractable for the subcategory of $\mathcal{C}$ induced by the image of $J$. For example, $J X$ is the apex of a cone over $G \circ \operatorname{cod}: J X \downarrow G \rightarrow \mathcal{C}$, so there exists a unique factorization through the limiting cone $\left(T X, \Psi_{X}\right), \eta_{X}: J X \rightarrow T X$.
(2) $\Phi$ is defined by taking $D$ to the relative algebra ( $G D, \chi$ ), where:

$$
(f: J X \rightarrow G D) \stackrel{\chi}{\longmapsto}\left(\Psi_{X}(f): T X \rightarrow G D\right)
$$

It remains to confirm the unit and multiplication laws are satisfied.

Clearly, the unit law holds as $\eta_{X}$ is formed as the unique factorization of the canonical cone with apex $J X$ through the limiting cone $\left(T X, \Psi_{X}\right)$, so:

$$
\Psi_{X}(f) \circ \eta_{X}=f
$$

Given $f: J X \rightarrow G D$ and $k: J Z \rightarrow T X$, we need to prove the multiplication law:

$$
\chi(\chi(f) \circ k)=\chi(f) \circ k^{\#}
$$

By definition, we have:

$$
\chi(\chi(f) \circ k)=\Psi_{Z}\left(\Psi_{X}(f) \circ k\right)
$$

But $k^{\#}$ is defined as the unique factorization of the cone $\Psi_{k}=\Psi_{Z}\left(\Psi_{X}(-) \circ k\right)$ through the limit $\left(T X, \Psi_{X}\right)$. This implies:

$$
\chi(f) \circ k^{\#}=\Psi_{Z}\left(\Psi_{X}(f) \circ k\right)
$$

which proves the multiplication law.
(3) Consider some $F{ }_{J} \dashv G$ and the corresponding natural isomorphism:

$$
\gamma_{X, D}: \mathcal{D}(F X, D) \cong \mathcal{C}(J X, G D)
$$

The component of the relative unit $\iota_{X}: J X \rightarrow G F X$ forms an initial object in the co-comma category $J X \downarrow G$, and so determines the limit up to isomorphism:

$$
\lim _{\leftarrow}(G \circ \operatorname{cod}) \cong G \circ \operatorname{cod}\left(\iota_{X}\right)
$$

Explicitly, $\left(G F X, G\left(\gamma^{-1}(-)\right)\right)$ forms a cone over $G \circ \operatorname{cod}: J X \downarrow G \rightarrow \mathcal{C}$, and $\Psi_{X}\left(\iota_{X}\right)$ is the unique factorization of the limit $T X$ through $G F X$. Therefore, $G F X \cong T X$.
$\Phi \circ F(X)$ is defined in (2) as the map:

$$
(h: J Y \rightarrow T X) \mapsto\left(\Psi_{Y}(h): T Y \rightarrow T X\right)
$$

By the above result, this is equivalent to $h \mapsto G\left(\gamma^{-1}(h)\right)$, which is the definition of $h^{\#}$, for the relative monad formed by the relative adjunction $F{ }_{J} \dashv G$.

In the classical case, we refer to this as the codensity monad, as $G$ is codense if and only if the unit of the codensity monad is an isomorphism. In the relative case, we obtain a similar correspondence, which we refer to as $J$-codensity.

Definition 2.30. $G$ is said to be $J$-codense if, for each $X$ :

$$
J X=\lim _{\rightleftarrows}(G \circ \operatorname{cod}: J X \downarrow G \rightarrow \mathcal{C})
$$

The comparison functor, $\Phi$, raises a more general question: When is a category equivalent to a category of relative (co) algebras? In the context of relative algebra, monadicity itself lacks a clear definition.

### 2.4 Monadicity: The Relative Case

Altenkirch et al. define a relative version of the Kleisli and Eilenberg-Moore categories for relative monads [1]. There is, however, no obvious analogue to Beck's Monadicity Theorem. There does not appear to be a way of classifying those categories which are equivalent to a category of algebras for a relative monad. One reason is that relative algebras do not carry an obvious free presentation, whereas classical algebras are presented as coequalizers of free algebras.

The following result, due to Eilenberg and Moore (see [7]), emphasizes the duality between monadicity and comonadicity, presented here as in (5):

Theorem 2.31. [5] Let $T$ be a monad on a category $\mathcal{X}$. If $T$ has a right adjoint $S$, the forgetful functor $U^{T}: \mathcal{X}^{T} \rightarrow \mathcal{X}$ is comonadic for a comonad $(S, \varepsilon, \delta)$.

The proof relies upon (co)free presentation of (co)algebras and makes heavy use of adjoint symmetry, neither of which are present in the relative case. Nevertheless, we are able to prove a relative version of Theorem 2.31 which, in the specific case of $\mathcal{N}$, gives a canonical presentation of internal presheaves as relative co-algebras (Theorem 4.24). We prove Theorem 2.33 by a series of three lemmas, making repeated use of two basic properties of relative adjunctions that are stated as a preliminary lemma.

Lemma 2.32. Given a relative adjoint $G \dashv_{J} F$, defined by the natural isomorphism:

$$
\Psi_{D, X}: \mathcal{C}(G D, J X) \cong \mathcal{D}(D, F X)
$$

1. $\Psi^{-1}=\tilde{\varepsilon} \circ G(-)$
2. $\Psi(h \circ G(f))=\Psi(h) \circ f$

Proof. The relative co-unit $\tilde{\varepsilon}=\Psi^{-1}\left(i d_{F}\right): G F \Rightarrow J$ satisfies the universal property that, given any morphism $f: G D \rightarrow J X$, there is a unique morphism $\Psi(f)$, such that:

$$
\tilde{\varepsilon}_{X} \circ G \Psi(f)=f
$$

Now, given any $g: D \rightarrow F X$, just take $f=\Psi^{-1}(g)$ and the first equivalence follows immediately. The second equivalence then follows:

$$
\begin{aligned}
& \Psi(h) \circ g \\
& \sim \Psi^{-1}(\Psi(h) \circ g) \\
& =\tilde{\varepsilon} \circ G(\Psi(h)) \circ G(g) \\
& =h \circ G(g) \\
& \sim \Psi(h \circ G(g))
\end{aligned}
$$

Theorem 2.33. Given a monad $\left(T, \eta^{T}, \mu\right), T: \mathcal{C} \rightarrow \mathcal{C}$, and a relative adjunction $T \dashv_{F} S$, defined by the natural isomorphism:

$$
\Psi: \mathcal{C}(T C, F D) \cong \mathcal{C}(C, S D)
$$

and relative co-unit $\varepsilon=\Psi^{-1}\left(i d_{S}\right): T S \Rightarrow F$. There is a canonical relative comonad $(S, \tilde{\varepsilon},())$ along $F$, with base functor $S$. The relative comonad arises from a co-freeforgetful relative adjunction $U \dashv_{F} G$ between $\mathcal{C}, \mathcal{D}$, and the category of $T$-Algebras, $\mathcal{C}^{T}$.


Proof. The proof corresponds to Lemmas 2.34, 2.35 and 2.36.
Lemma 2.34. The functor $G: \mathcal{D} \rightarrow \mathcal{C}^{T}$, defined by the action on objects:

$$
D \mapsto\left(S D, \omega_{D}\right)
$$

where

$$
\omega_{D}=\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right)
$$

and whose action on morphisms is induced by the functor $S$, maps each $D \in \mathcal{D}$ to a "co-free" T-algebra.

Proof. First, we need to prove that $\left(S D, \omega_{D}\right)$ satisfies the unit and multiplication properties of a $T$-algebra.

The unit law corresponds to the commutative diagram


$$
\begin{aligned}
& \omega_{D} \circ \eta_{S D}^{T}=\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \eta_{S D}^{T} \\
& \sim \Psi^{-1}\left(\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \eta_{S D}^{T}\right) \\
& =\varepsilon_{D} T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \eta_{S D}^{T} \\
& =\varepsilon_{D} \circ \mu_{S D} \circ T \eta_{S D}^{T}=\varepsilon_{D} \\
& =\Psi^{-1}\left(i d_{S D}\right)
\end{aligned}
$$

$$
\sim \Psi^{-1}\left(\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \eta_{S D}^{T}\right) \quad \text { (relative transpose) }
$$

Hence, $\left(S D, \omega_{D}\right)$ satisfies the unit law of a $T$-algebra.

For the multiplication law, we need to prove:

$$
\begin{array}{rlr}
\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \mu_{S D}=\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) & \\
& \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \mu_{S D} & \\
& \sim \Psi^{-1}\left(\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \mu_{S D}\right) & \\
& =\varepsilon_{D} \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \mu_{S D} & \\
& =\varepsilon_{D} \circ \mu_{S D} \circ T \mu_{S D} & \\
& =\varepsilon_{D} \circ \mu_{S D} \circ \mu_{T S D} & \text { (transpose) } \\
& =\varepsilon_{D} \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \mu_{T S D} & \\
& =\varepsilon_{D} \circ \mu_{S D} \circ T^{2} \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) & \\
& =\varepsilon_{D} \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T^{2} \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) & \\
& =\Psi^{-1}\left(\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right)\right) & \\
& \sim \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) & \text { (traturality of } \mu) \\
\hline \text { (transpose) }
\end{array}
$$

Therefore, $G D=\left(S D, \omega_{D}\right)$ is a $T$-algebra.

The action of $G$ on morphisms is inherited from $S$, thus we need only prove that, given $g: D \rightarrow E, S(g)$ is a morphism of $T$-algebras $\left(S D, \omega_{D}\right) \rightarrow\left(S E, \omega_{E}\right)$. In other words, we wish to prove:

\[

\]

Lemma 2.35. $G$ is an $F$-relative right adjoint to the forgetful functor $U: \mathcal{C}^{T} \rightarrow \mathcal{C}$, with relative co-unit $\tilde{\varepsilon}: U G \Rightarrow F$.

Proof. Define the relative co-unit $\tilde{\varepsilon}: U G \Rightarrow F$ as $\varepsilon \circ \eta_{S}^{T}$. Furthermore, given some $T$ algebra $(k, C)$ and an object $D \in \mathcal{D}$, we define the natural isomorphism corresponding to $U \dashv_{F} G$ as ${ }^{19}$

$$
\hat{\Psi}_{k, D} \equiv \Psi(-\circ k): \mathcal{C}(U(k, C), F D) \rightarrow \mathcal{C}^{T}\left((k, C),\left(S D, \omega_{D}\right)\right)
$$

Given a morphism $f: C \rightarrow F D$, we obtain the factorization:


[^48]We first prove that $\Psi(f \circ k)$ is a morphism of $T$-algebras:

$$
\begin{array}{lr}
\omega_{D} \circ T \Psi(f \circ k)=\Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T \Psi(f \circ k) & \\
\sim \varepsilon_{D} \circ \mu_{S D} \circ T^{2} \Psi(f \circ k) & \text { (transpose) }  \tag{transpose}\\
=\varepsilon_{D} \circ T \Psi(f \circ k) \circ \mu_{T S D} & \text { (naturality of } \mu \text { ) } \\
=f \circ k \circ \mu_{T S D}=f \circ k \circ T k & \text { (multiplication property of } T \text {-algebras) } \\
=\Psi^{-1}(\Psi(f \circ k) \circ k) & \text { (Lemma 2.32) } \\
\sim \Psi(f \circ k) \circ k & \text { (transpose) }
\end{array}
$$

It remains to prove $f$ factors uniquely through $\tilde{\varepsilon}_{D}$ :

$$
\begin{array}{lr}
\varepsilon_{D} \circ \eta_{S D}^{T} \circ \Psi(f \circ k)=\varepsilon_{D} \circ T \Psi(f \circ k) \circ \eta_{C}^{T} & \text { (naturality) } \\
=f \circ k \circ \eta_{C}^{T} & \text { (Lemma 2.32) } \\
=f & \text { (by unit law of } T \text {-algebra } k \text { ) }
\end{array}
$$

To show uniqueness, consider some $h: C \rightarrow S D$ such that:

$$
\varepsilon_{D} \circ \eta_{S D}^{T} \circ h=f
$$

we obtain the following chain of equivalences, proving $\Psi(f \circ k)=h$ :

$$
\begin{array}{ll}
\Psi(f \circ k)=\Psi\left(\varepsilon_{D} \circ \eta_{S D}^{T} \circ h \circ k\right) \\
=\Psi\left(\varepsilon_{D} \circ \eta_{S D}^{T} \circ \omega_{D} \circ T h\right) & \\
=\Psi\left(\varepsilon_{D} \circ \eta_{S D}^{T} \circ \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ T h\right) & \text { (as } h \text { is a morphism of } T \text {-algebras) } \\
=\Psi\left(\varepsilon_{D} \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \eta_{S D}^{T} \circ T h\right)=\Psi\left(\varepsilon_{D} \circ \mu_{S D} \circ \eta_{S D}^{T} \circ T h\right) \quad \text { (naturality) } \\
=\Psi\left(\varepsilon_{D} \circ T h\right)=h & \text { (unit law of monads, Lemma 2.32) }
\end{array}
$$

Therefore $U \dashv_{F} G$.
Lemma 2.36. We construct the relative comonad (S, $\tilde{\varepsilon}, \hat{( })$ ).

Proof. We have already constructed $S$ and $\tilde{\varepsilon}$ :

$$
\begin{aligned}
& S=U \circ G \\
& \tilde{\varepsilon}=\varepsilon \circ \eta_{S}^{T}: S \rightarrow F
\end{aligned}
$$

All that remains is to define $\hat{( })$ :

$$
\hat{()} \equiv \Psi(-\circ \omega):|S \downarrow F| \rightarrow|S \downarrow S|
$$

We claim:

$$
\left(S, \tilde{\varepsilon},(\hat{)}) \equiv\left(U G, \varepsilon \circ \eta_{T}^{S}, \Psi\left(-\circ \Psi\left(\varepsilon \circ \mu_{S}\right)\right)\right)\right.
$$

satisfies the properties defining a relative comonad.

1) $\hat{\tilde{\varepsilon}}=i d$

$$
\begin{aligned}
& \hat{\tilde{\varepsilon}}_{D}=\Psi\left(\varepsilon_{D} \circ \eta_{S D}^{T} \circ \omega_{D}\right) \\
& =\Psi\left(\varepsilon_{D} \circ T \Psi\left(\varepsilon_{D} \circ \mu_{S D}\right) \circ \eta_{T S D}^{T}\right) \quad \text { (definition of } \omega_{D}, \text { naturality of } \eta^{T} \text { ) } \\
& =\Psi\left(\varepsilon_{D} \circ \mu_{S D} \circ \eta_{T S D}^{T}\right)=\Psi\left(\varepsilon_{D}\right)=\Psi\left(\Psi^{-1}\left(i d_{D}\right)\right)=i d_{D} \quad \text { (Lemma 2.27, unit law) }
\end{aligned}
$$

2) $\tilde{\varepsilon}_{D} \circ \hat{k}=k$

$$
\begin{array}{lr}
\tilde{\varepsilon}_{D} \circ \hat{k}=\varepsilon_{D} \circ \eta_{S D}^{T} \circ \Psi\left(k \circ \omega_{C}\right) & \\
=\varepsilon_{D} \circ T \Psi\left(k \circ \omega_{C}\right) \circ \eta_{S C}^{T} & \text { (naturality of } \eta^{T} \text { ) } \\
=k \circ \omega_{C} \circ \eta_{S C}^{T}=k & \text { (unit law of } T \text {-algebras) }
\end{array}
$$

3) $\hat{h} \circ \hat{k}=\widehat{h \circ \hat{k}}$

$$
\begin{array}{ll}
k: S C \rightarrow F D, h: S D \rightarrow F E & \\
\hat{h} \circ \hat{k}=\Psi\left(h \circ \omega_{D}\right) \circ \Psi\left(k \circ \omega_{C}\right) & \\
=\Psi\left(h \circ \omega_{D} \circ T \Psi\left(k \circ \omega_{C}\right)\right) & \\
=\Psi\left(h \circ \Psi\left(k \circ \omega_{C}\right) \circ \omega_{C}\right) & \text { (Lemma 2.32) } \\
=\widehat{h \circ \hat{k}} &
\end{array}
$$

Following the construction of a relative comonad $\left(S, \tilde{\varepsilon},(\hat{)})\right.$ from a monad $\left(T, \eta^{T}, \mu\right)$, we investigate the relationship between the categories of $T$-algebras and $S$-relative coalgebras.

Proposition 2.37. $\lambda(k, C) . \Psi(-\circ k)$ forms a full and faithful functor from the category of $T$-algebras to the category of relative $S$-co-algebras.

Proof. Given a $T$-algebra $(k, C)$, we claim $\tau_{k}=\Psi(-\circ k)$ defines a relative co-algebra at $C$. Verification of the relative co-algebra conditions follow from similar methods used to verify the isomorphism $\hat{\Psi}$ defining the relative adjunction $U \dashv_{F} G$.
(1) For any $f: C \rightarrow F D, \tilde{\varepsilon} \circ \tau_{k}(f)=f$ :

$$
\begin{array}{lr}
\tilde{\varepsilon}_{D} \circ \tau_{k}(f)=\tilde{\varepsilon}_{D} \circ \Psi(f \circ k) \\
=\varepsilon_{D} \circ \eta_{S D}^{T} \circ \Psi(f \circ k)=\varepsilon_{D} \circ T \Psi(f \circ k) \circ \eta_{C}^{T} & \text { (naturality) } \\
=f \circ k \circ \eta_{C}^{T}=f & \text { (unit law of } T \text {-algebras) }
\end{array}
$$

(2) Given $g: S D \rightarrow F E$ and $f: C \rightarrow F D, \tau_{k}\left(g \circ \tau_{k}(f)\right)=\hat{g} \circ \tau_{k}(f)$

$$
\begin{array}{lr}
\tau(g \circ \tau(f))=\Psi(g \circ \Psi(f \circ k) \circ k) & \\
=\Psi\left(g \circ \omega_{D} \circ T \Psi(f \circ k)\right) & \text { (morphism of } T \text {-algebras) } \\
=\Psi\left(g \circ \omega_{D}\right) \circ \Psi(f \circ k) &  \tag{Lemma2.32}\\
=\hat{g} \circ \tau_{k}(f) &
\end{array}
$$

Given a map of $T$-algebras $f:(k, C) \rightarrow(j, D), f$ is also a map of $S$-relative co-algebras $f:\left(C, \tau_{k}\right) \rightarrow\left(D, \tau_{j}\right)$.

$$
\begin{array}{lr}
\tau_{k}(g \circ f)=\Psi(g \circ f \circ k) & \\
\sim g \circ f \circ k & \text { (transpose) } \\
=g \circ j \circ T f & \text { (morphism of } T \text {-algebras) } \\
\sim \Psi(g \circ j \circ T f) & \text { (transpose) }  \tag{transpose}\\
=\Psi(g \circ j) \circ f=\tau_{j}(g) \circ f & \text { (Lemma 2.32) }
\end{array}
$$

As morphisms of algebras (resp. relative co-algebras) are simply morphisms of $\mathcal{C}$ satisfying what we have just shown to be equivalent conditions, the functor is full and faithful.

It is by no means clear that, as in the classical case, this should form an equivalence of categories. Even the property that $\lambda(k, C) \cdot \Psi(-\circ k)$ is a proper embedding appears to
require further assumptions. For example, one sufficient condition would be that each family of morphisms $|C \downarrow F|$ is jointly monic. Consider two distinct $T$-algebras ( $j, C$ ) and $(k, C)$, which form the same relative co-algebra. $\tau_{j}=\tau_{k}$ is equivalent to saying that for any $g: C \rightarrow F D$ :

$$
\Psi(g \circ j)=\Psi(g \circ k)
$$

As $\Psi$ forms a (natural) isomorphism between hom-sets, this implies:

$$
\tau_{j}=\tau_{k} \sim \forall g \in|C \downarrow F| . g \circ j=g \circ k
$$

Hence, if the resulting family of all such maps is jointly monic (indeed, if any individual arrow is monic), it must be the case that $(j, C)=(k, C)$.

The following results consider two special cases.

Corollary 2.38. $\lambda(k, C) . \Psi(-\circ k)$ restricts to a proper embedding of the subcategory of $T$-algebras, whose objects are in the image of $F$, into the category of $S$-relative coalgebras.

Proof. If $(a, F C)$ and $(b, F C)$ are mapped to the same relative co-algebra, $\tau_{a}\left(i d_{F C}\right)=$ $\tau_{b}\left(i d_{F C}\right)$. Hence $\Psi(a)=\Psi(b)$, so $a=b$.

Corollary 2.39. If $\mathcal{C}$ has an object $C$, "isolated" from $F$, in the sense that $|C \downarrow F|$ is empty, any $T$-algebra at $C$ will be mapped to the same (trivial) relative co-algebra. In the case that there are multiple $T$-algebras at $C$, it follows that $\lambda(k, C) . \Psi(-\circ k)$ is not a proper embedding.

These two results speak to earlier observations regarding relative adjunctions and relative (co)monads [68, 1].

1. It is particularly convenient for the relative functor $F$ to be (co)dense in $\mathcal{C}$. In this case, we are able to eliminate the possibility of "isolated" objects in $\mathcal{C}$, the presence of which greatly confounds the generalization of monads to relative monads.
2. The equivalence of the Manes-style presentation of (co)algebras to the classical presentation can be reduced to the fact that a Manes-style (co)monad is entirely determined by its action on the identity morphism. Therefore, in the case that an identity morphism exists as an object in the comma category $C \downarrow F$ (i.e. when $C$ is in the image of $F), \lambda(k, C) . \Psi(-\circ k)$ is injective on objects. An intuitive interpretation of relative (co)algebras is that they form a universal object in the associated comma category in the same way that the identity morphism $1_{C}$ is an initial object in the co-slice category $C / \mathcal{C}$. One can, by the pointwise construction of Kan extensions, recover from this fact some intuition for an observation in [1] that relative monads permitting (left) Kan extensions are well-behaved.

Proposition 2.40. Given an arbitrary $F$-relative comonad $(S, \tilde{\varepsilon},())$, where $F$ is full and faithful and $\tilde{\varepsilon}$ is monic, for any $S$-relative co-algebra $(\tau, F D), \tau$ is uniquely determined by $\tau\left(i d_{F D}\right)$. For any $F f: F D \rightarrow F E$ :

$$
\tau(F(f))=S(f) \circ \tau\left(i d_{F D}\right)
$$

Proof. We use the property of $\tilde{\varepsilon}_{E} \circ \tau(g)=g$ and show $\tilde{\varepsilon}_{E} \circ S(f) \circ \tau\left(i d_{F D}\right)=F(f)$. The result follows from the assumption that $\tilde{\varepsilon}$ is monic.

$$
\begin{aligned}
& \tilde{\varepsilon}_{E} \circ S(f) \circ \tau\left(i d_{F D}\right) \\
& =F(f) \circ \tilde{\varepsilon}_{D} \circ \tau\left(i d_{F D}\right)
\end{aligned}
$$

$$
=F(f) \quad(\text { as } \tau \text { is a relative co-algebra })
$$

In the standard case, where $\tau(f)$ is $S(f) \circ j$, for some (standard) co-algebra $j$, the above result holds with no assumptions on $\tilde{\varepsilon}$.

## Chapter 3

## Categories of Stratified Sets

Much of the intuition for categorical abstraction comes from instantiation in Set. If one is to consider NF as a foundation for category theory, one needs to consider not only the categorical properties of $\mathcal{N}$, but how intuition derived from $\mathcal{N}$ extends to more general categories ${ }^{1}$

## Motivation: Displaying Indexed Families

Example 3.1 (Indexed families in ZF). In the category of ZF sets, $S e t / I$ and $S e t^{I}$, the category of functions with codomain $I$ and the category of functors from the discrete category $I$ to $S e t$, form an equivalence of categories:

$$
\begin{gathered}
((f: C \rightarrow I) \in S e t / I) \mapsto\left(F: I \rightarrow \text { Set }: i \mapsto f^{-1}(i)\right) \\
\left((F: I \rightarrow S e t) \in S e t^{I}\right) \mapsto\left(f: \coprod_{i \in I} F(i) \rightarrow I: j \in F(i) \mapsto i\right)
\end{gathered}
$$

[^49]The important feature is: any $I$-indexed family can be represented internal to Set, as a morphism whose fibres correspond to the members of the indexed family. Therefore, we view $\mathcal{C} / I$ as the "correct" notion of $I$-indexed families in an arbitrary category $\mathcal{C}$.

Example 3.2 (Indexed families in NF). In the category of NF sets, the equivalence is not between $\mathcal{N} / I$ and $\mathcal{N}^{I}$, but between $\mathcal{N} / I$ and $\mathcal{N}^{T I}$. Given a function $f: C \rightarrow I$ in NF , the inverse $f^{-1}$ is not generally defined by a stratified formula, as it is inhomogeneous. $f^{-1}$ can, however, be defined as a function $T I \rightarrow P C$, between the set of singletons of elements of $I$ and the powerset of $C$. Thus, one direction of the correspondence stated above is now:

$$
((f: C \rightarrow I) \in \mathcal{N} / I) \mapsto\left(F: T I \rightarrow \mathcal{N}:\{i\} \mapsto f^{-1}(i)\right)
$$

On the other hand, $\coprod_{i \in I} F(\{i\})$ and $F(i)$ are type-level. Thus, the function $f: \coprod_{i \in I} F(\{i\}) \rightarrow$ $I: j \in F(\{i\}) \mapsto i$ is unstratified, as $j$ is a type below $i$. To "lower" the type of $F$ to that of $f$, we need to be able to reduce the type of $I$. This can be accomplished in the following special case:

$$
\left((F: T I \rightarrow \mathcal{N}) \in \mathcal{N}^{T I}\right) \mapsto\left(f: \coprod_{i \in I} F(\{i\}) \rightarrow I: j \in F(\{i\}) \mapsto i\right)
$$

This does not mean that NF has only families indexed by sets the same size as a set of singletons ${ }^{2}$ Rather, $\mathcal{N}$ can only display families that are indexed by sets of singletons ${ }^{3}$ Indeed, one possible characterization of $\mathcal{N}$ is as a category with display maps for exactly those families indexed by sets in the image of a given endofunctor $T$.

This chapter investigates the general properties of categories of stratified sets, focusing on NF and KF in particular ${ }_{4}^{4}$ KF is equivalent to Mac Lane set theory, restricted to stratified $\Delta_{0}$-separation. It is a convenient theory for us to study, as it is a subtheory of both ZF and NF. Our primary interest, however, remains $\mathcal{N}$.

A theme which runs over the course of this chapter is the existence of $T$-relative adjoint relationships in stratified theories, where we would expect adjunctions. The relationship

[^50]between $T$-relative structures and typed syntax is apparent in a couple of ways. First, while adjoint symmetry is lacking in the relative case, NF appears agnostic to whether an adjoint is exchanged for a $T$-right or $T$-left relative adjoint. In every case we have found, not only are both available, but they are coherent in the sense that they paste to form a symmetric lift. Second, we can view these $T$-relative structures as "term structures" - satisfying universal properties in a way that feels distinctly syntactic.

The product-exponential adjunction, $A \times-\dashv(-)^{A}$, gives rise to a relative adjoint structure, referred to as modified-cartesian closure. The stability of modified-cartesian closure under localization informs the development of an "NF-topos" and a stratified analogue to the fundamental theorem of topos theory. Finally, we consider some ideas of "smallness" in NF. Given the presence of big sets like $V, \mathcal{N}$ is a compelling candidate for a class category, with a cartesian closed subcategory of small classes (i.e. sets). We can prove the folklore definition of small sets in NF - the strongly cantorian sets - forms a subcategory of NF that is a topos. Extending this to the category theoretic (fibre-wise) notion of smallness, however, requires extending NF by an additional axiom $S C U$. In this extension we are able to obtain some results of set theoretic interest.

### 3.1 Defining a Category of Stratified Sets

### 3.1.1 Set Theories of Interest

Our primary focus will be on the category of NF sets but, where possible, we seek to prove results in (the weaker theory) KF. This is a useful exercise for us in two ways:

1. As $N F=K F+\exists y \forall x . x \in y$, we gain a better understanding of what category theoretic properties are, resepectively, permitted and restricted by the existence of a universal set 5
2. Both Mac and NF are extensions of KF. Thus, by studying KF, we gain insight

[^51]into a common restriction of both $\mathcal{N}$ and a set theory corresponding to the internal language of an elementary topos.

In Section 1.4, we introduced three extenstions of KF. The first extension, KFI (KF + Inf), is required for the implementation of Quine Pairs. Infinity is a theorem of NF, but the following pair of axioms are not $\sqrt{6}^{6}$

1. IO: The principle that every set is the same size as a set of singletons.
2. CE: The principle that every family of pairwise disjoint sets is the same size as a set of singletons.
$\mathrm{KF}+\mathbf{I O}$ is not as strong as the assumption that all sets are cantorian, much less strongly cantorian - assuming the latter would simply give Mac. But IO is still inconsistent with the existence of a universal set. CE is of interest to both KF and NF, and is connected to the existence of general (internal) colimits in $\mathcal{N}$.

Lemma 3.3. In $K F+\boldsymbol{C E}$, any partition of a set $x$ can be assigned a representation, which will be the same type as $x$ in a valid stratification.

### 3.1.2 The Basic Category

It is not difficult to see that any model of KF or NF is a category. The objects are just the sets of the model and the morphisms consist of functional relations (that are also sets of the model) tagged with a corresponding codomain.

Definition 3.4. The category of NF sets, $\mathcal{N}$, is defined as follows:

Objects: In general, the objects are just those of a generic model of NF.7 In NF, the collection of objects, denoted $|\mathcal{N}|$, is simply the universal set $V$.

[^52]Morphisms: $\operatorname{Mor}(\mathcal{N})$ consists of whatever functional relations are sets of the model, tagged with a given codomain. In NF we can form the set of all morphisms explicitly, where Fun denotes the set of all functional relations:

$$
\operatorname{Mor}(\mathcal{N})=\{\langle f, y\rangle \mid f \in \operatorname{Fun} \wedge \forall w \cdot(\exists x .\langle x, w\rangle \in f) \Longrightarrow w \in y\}
$$

Composition and Identity: Composition of functional relations is clearly homogeneous, and is just inherited from composition in the model. Similarly it is clear that, given any set $x$, we are able to define the functional relation $\delta_{x}=\{\langle y, y\rangle \mid y \in x\}$. The identity morphism is then uniquely determined as the pair $\left\langle\delta_{x}, x\right\rangle$.

Furthermore, in NF, both composition and the function assigning identity are homogeneous and definable by stratified comprehension.

$$
\begin{align*}
& \operatorname{dom}: \operatorname{Mor}(\mathcal{N}) \rightarrow V:\langle f, y\rangle \mapsto \pi_{0} " f  \tag{Source}\\
& \operatorname{cod}: \operatorname{Mor}(\mathcal{N}) \rightarrow V:\langle f, y\rangle \mapsto y  \tag{Target}\\
& \operatorname{Mor}(\mathcal{N}) \times_{\text {dom }} \operatorname{Mor}(\mathcal{N})=\left\{\langle\langle f, y\rangle,\langle g, z\rangle\rangle \mid \pi_{0} " g=y\right\} \\
& m: \operatorname{Mor}(\mathcal{N}) \times_{\text {dom }} \operatorname{Mor}(\mathcal{N}) \rightarrow \operatorname{Mor}(\mathcal{N}):\langle\langle f, y\rangle,\langle g, z\rangle\rangle \mapsto\langle g \circ f, z\rangle  \tag{Composition}\\
& i: V \rightarrow \operatorname{Mor}(\mathcal{N}): x \mapsto\left\langle\delta_{x}, x\right\rangle \tag{Identity}
\end{align*}
$$

Remark ( $\mathcal{N}$ Internalizes Itself). A category theorist will already have noted that, as all the objects and morphisms in Definition 3.4 are sets of NF, we have just defined an internal category in the category of NF sets ${ }^{8}$ Paradigmatically, $\mathcal{N} \in \operatorname{cat}(\mathcal{N})$ - for that matter, $\operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\mathcal{N})($ just as $V \in V) \cdot 9$

Despite the fact that $\mathcal{N}$ internalizes itself, however, we do not eliminate the need to consider certain important categories that are external (i.e. that are not internal to $\mathcal{N}$ ).

Example 3.5 (A Locally (NF-)Small Category). The category of strongly cantorian sets, $S C$, is a full subcategory of $\mathcal{N}$. While the collection of maps between any two strongly cantorian sets is a set, $|S C| \in V$ implies the Burali-Forti paradox ${ }^{100}$ Therefore,

[^53]as $|S C|$ is a class, $S C$ is locally small (in the category theoretic sense). Furthermore, the hom-set between any two strongly cantorian sets is itself strongly cantorian, thus $S C$ is also locally $N F$-small ${ }^{11}$

In fact, while we will occasionally consider another external category in passing, any category considered in this thesis (in the context of NF) will have hom-sets that are truly sets of NF. Therefore, the only interesting form of local smallness we will need to consider is local NF-smallness.

Remark (Limits to Externality). Not only will each external category we consider have hom-sets that are genuinely sets of our model. As every ML class is a subclass of the universal set, each external category will be a subcategory of $\mathcal{N}$.

## Functors

One reason for having brought up internal categories ahead of schedule is that they provide a useful setting for understanding the functors that appear naturally in our development of $\mathcal{N}$.

A functor $F$ is defined as a pair of (possibly external) functions, $F_{0}$ and $F_{1}$, acting on objects and morphisms, respectively. $F_{0}$ and $F_{1}$ must be coherent in the classical sense:

$$
\begin{array}{lr}
f: a \rightarrow b \mapsto F_{1}(f): F_{0}(a) \rightarrow F_{0}(b) & \text { (Coherence between } \left.F_{0} \text { and } F_{1}\right) \\
F_{1}\left(i d_{a}\right)=i d_{F_{0}(a)} & \text { (Preservation of Identity) } \\
F_{1}(g \circ f)=F_{1}(g) \circ F_{1}(f) & \text { (Distributive Property) }
\end{array}
$$

If $F_{0}$ and $F_{1}$ are sets, then $F$ is just an internal functor in $\mathcal{N}$.
Example 3.6 (Internal Functor). The identity functor, $I: \mathcal{N} \rightarrow \mathcal{N}$, is defined by sets:

$$
\begin{aligned}
& I_{0}=\{\langle x, x\rangle \mid x \in \operatorname{Obj}(\mathcal{N})\} \\
& I_{1}=\{\langle f, f\rangle \mid f \in \operatorname{Mor}(\mathcal{N})\}
\end{aligned}
$$

[^54]However, many functors required for the development of category theory in $\mathrm{NF}-T$ being the obvious example - have action maps that are external (i.e. are class maps).

Example 3.7 ((Definable) External Functor). $T$ is defined by inhomogeneous operations:

$$
\begin{aligned}
& T_{0}: x \mapsto \iota " x \\
& T_{1}: f: x \rightarrow y \mapsto \operatorname{RUSC}(f): \iota " x \rightarrow \iota " y
\end{aligned}
$$

While neither $T_{0}$ or $T_{1}$ are definable as sets, we have shown that there are stratified formulae $\phi_{0}$ and $\phi_{1}$, such that:

$$
\forall x \exists y \forall z . z \in y \Longleftrightarrow \phi_{i}(z, x) . \sim . \forall x \exists y . y=T_{i}(x)
$$

## External Functors of $\mathrm{NF}=$ Lateral Functors in ML

In every case encountered within the context of this thesis, external functors will be definable (albeit inhomogeneous), in the sense that there exist stratified formulae $\Psi_{0}$ and $\Psi_{1}$ such that $F_{0}$ and $F_{1}$ correspond to the action $F_{i}(x)=\left\{y \mid \Psi_{i}(y, x)\right\} .{ }^{12}$ But there is no reason to think - in fact it is extremely unlikely - that all external functors will be definable $\sqrt{13}$ Therefore, we need to develop a semantic conception of external functors. Here we can not avoid engaging with the meta-theory (i.e. the theory of classes in which the actions maps of an external functor exist).

As we wish to make as few assumptions as possible regarding the ambient theory within which our model of NF exists, we choose to work in ML ${ }^{[14}$ Thus, to rigorously define external functors, we return to our development of lateral functions in Section $1.4{ }^{15}$

[^55]Recall that a class function in ML is lateral if there exists some $n \in N$ such that $\operatorname{set}\left(F \mid I_{n}^{-1}\right)$ or $\operatorname{set}\left(I_{n} \circ G\right)$, where $I$ is the class function corresponding to the action: $\langle n, x\rangle \mapsto \iota^{n \iota} x .^{16}$ As such, $F$ is a lateral functor in ML if its respective action maps $F_{0}$ and $F_{1}$ are lateral functions.

Definition 3.8. A functor $F: \mathcal{C} \rightarrow \mathcal{D}$ in ML is n-lateral when both $F_{0}: C_{0} \rightarrow D_{0}$ and $F_{1}: C_{1} \rightarrow D_{1}$ are $n$-lateral functions. In the context of NF, we take lateral functors to be the appropriate definition of external functors between categories.

First, we should note a consequence of Definition 3.8: the collection of 0-lateral functors between any pair of $\mathcal{N}$-internal categories $\mathbb{C}$ and $\mathbb{D}$ is just the (set of) objects for the internal functor category, $\operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}]{ }^{[17}$ Thus, given a pair of internal categories, $\mathbb{C}$ and $\mathbb{D}$, the internal functor category forms a proper subcategory of the full (i.e. external) functor category. We therefore note the following convention, given a pair of (internal) $\mathbb{C}$ and $\mathbb{D}, 118$

$$
\begin{aligned}
& \operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}] \\
& {[\mathbb{C}, \mathbb{D}]}
\end{aligned}
$$

(Internal Functor Category)
(External Functor Category)
Remark. Unlike ZF(C), classes are not determined by (cardinal) size in NF. Therefore, as is the case with $T$, in the above example, one can have external functors between internal categories in $\mathcal{N}$. This implies a property that seems unique to stratified set theories: Given any pair of small categories, the corresponding external and internal functor categories need not be equivalent: $\sqrt{19}$

$$
\exists \mathbb{C}, \mathbb{D} \in \operatorname{cat}(\mathcal{N}) \cdot \operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}] \subsetneq[\mathbb{C}, \mathbb{D}]
$$

someone with no previous background in NF, this may even be advisable. That being said, if the reader is interested in issues relating to a truly closed foundation for category theory, the details should still be of interest.
${ }^{16}$ See Definitions 1.40 and 1.41
${ }^{17}$ Our use of "internal" is a bit of pun. The functors of $\operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}]$ are, of course, internal. But so is $\operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}]$.
${ }^{18}$ For the reader uncomfortable with the idea of thinking about internal category theory prior to its formal introduction in Section 4.1, we would note that $\mathcal{N}$ is itself an internal category. Thus, the reader is free to think of the internal/external distinction in the specific case of endofunctors, $\mathcal{N} \rightarrow \mathcal{N}$.
${ }^{19}$ Some prominent examples are $F a m, \mathcal{Y}: \operatorname{cat}(\mathcal{N}) \rightarrow \operatorname{cat}(\mathcal{N})$. Both are in $[\operatorname{cat}(\mathcal{N}), \operatorname{cat}(\mathcal{N})]$ and not in the internal functor category, $\operatorname{cat}(\mathcal{N})[\operatorname{cat}(\mathcal{N}), \operatorname{cat}(\mathcal{N})]$.

But this raises an important question: Is the (external) category of all functors between $\mathbb{C}$ and $\mathbb{D}$ a category in $M L$ ? ${ }^{20}$

On its face it need not be - an ML class contains only sets. Indeed, if we were to consider all possible pairs of ML class functions, satisfying functoriality conditions, this would be problematic. But Definition 3.8 allows us to (partially) evade this complication.

In Lemma 3.11, we use the notation $T \mathbb{C}$ to denote the (internal) category whose set of objects, set of morphisms, and structure maps are precisely ' $T$ of' those defining $\mathbb{C}$. As $T: \mathcal{N} \rightarrow \mathcal{N}$ is full and faithful, so is the induced endofunctor on $\operatorname{cat}(\mathcal{N})$.

Given any $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, we can define a unique category $T \mathbb{C} \in \operatorname{cat}(\mathcal{N})$ as above. But $T \notin$ $\operatorname{cat}(\mathcal{N})[\operatorname{cat}(\mathcal{N}), \operatorname{cat}(\mathcal{N})]-T$ is not a $\mathcal{N}$-internal functor. Instead, $T \in[\operatorname{cat}(\mathcal{N}), \operatorname{cat}(\mathcal{N})]$ - $T$ is a 1-lateral functor in ML. Importantly, $T$ is an isomorphic functor in ML. The primary distinction between working in $\mathcal{N}$ and in the category of ML classes is: In general, $T \mathbb{C}$ and $\mathbb{C}$ are not isomorphic in cat $(\mathcal{N})$, despite being (externally) isomorphic in $M L$.

As we ubiquitously employ $T$ to denote the $T$-functor of $\mathcal{N}$, we denote the functor(s) induced by $I_{n}$, in ML, $\tilde{I}_{n}$.

Definition 3.9. Recall, the ML class function $I$, defined by the action $\langle n, x\rangle \mapsto \iota^{n \iota} x$, and the assiociated class functions $I_{n}$, defined by the action $x \mapsto \iota^{n ‘} x$. Each $I_{n}$ induces an $n$-lateral (isomorphic) functor $\tilde{I}_{n}: \mathcal{N} \rightarrow \mathcal{N}$, defined by the actions:

$$
\begin{aligned}
& \left(\tilde{I}_{n}\right)_{0}=I_{n}: V \rightarrow V: x \mapsto \iota^{n \iota} x \quad \text { (Action on } \\
& \left(\tilde{I}_{n}\right)_{1}=R U S C^{n}: \operatorname{Mor}(\mathcal{N}) \rightarrow \operatorname{Mor}(\mathcal{N}):\langle f, y\rangle \mapsto\left\langle R U S C^{n \iota} f, I_{n}{ }^{‘} y\right\rangle
\end{aligned}
$$

(Action on Morphisms)
Given any (possibly external) subcategory $\mathcal{C}$ of $\mathcal{N}, \tilde{I}_{n}$ restricts to an $n$-lateral isomorphism, $\mathcal{C} \cong T^{n} \mathcal{C}$.

[^56]Example 3.10. From Definition 3.9, we see the composite:

$$
i \circ \tilde{I}_{1}: \mathcal{N} \rightarrow T \mathcal{N} \hookrightarrow \mathcal{N}
$$

is just the standard $T$-functor, $T: \mathcal{N} \rightarrow \mathcal{N}$.

But we should note that the composing $\tilde{I}_{1}$ with inclusion, $T \mathcal{N} \hookrightarrow \mathcal{N}$, has a nontrivial impact, even in ML. Indeed, the inverse of $\tilde{I}_{1}, \tilde{I_{-1}}: T \mathcal{N} \rightarrow \mathcal{N}$, cannot be extended to an endofunctor on $\mathcal{N}$. Thus, while $\tilde{I}_{1}$ witnesses $\mathcal{N} \cong T \mathcal{N}$ in ML, the $T$-functor remains a full and faithful embedding but, as in NF, does not witness an equivalence of categories ${ }^{21}$

Lemma 3.11. Given categories $\mathbb{C}, \mathbb{D} \in \operatorname{cat}(\mathcal{N})$ a functor $F: \mathbb{C} \rightarrow \mathbb{D}$ in the ambient model of $M L$ is $n$-lateral if and only if $\tilde{F} \equiv F \mid I_{n}^{-1}: T^{n} \mathbb{C} \rightarrow \mathbb{D}$ is an internal functor in $\operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right]$. Likewise, $G: \mathbb{C} \rightarrow \mathbb{D}$ is-n-lateral if and only if $\tilde{G} \equiv I_{n} \circ G: \mathbb{C} \rightarrow T^{n} \mathbb{D}$ is an internal functor.

Proof. By definition, the action maps $F_{0}: C_{0} \rightarrow D_{0}$ and $F_{1}: C_{1} \rightarrow D_{1}$, of an $n$-lateral functor $F: \mathbb{C} \rightarrow \mathbb{D}$, are $n$-lateral functions. In other words:

$$
\text { For } i \in\{0,1\}, \operatorname{set}\left(\left\{\left\langle\iota^{n ‘} x, F_{i}{ }^{‘} x\right\rangle \mid x \in C_{i}\right\}\right) .
$$

Meanwhile,

$$
\text { For } i \in\{0,1\}, \iota^{n \iota} x \in T^{n} C_{i} \Longleftrightarrow x \in C_{i} .
$$

Thus, the action maps of $\tilde{F}$ are simply morphisms in $\mathcal{N}$. As the functoriality conditions are just inherited from those of $F, \tilde{F}$ is just an internal functor $\tilde{F}: T \mathbb{C} \rightarrow \mathbb{D}$. The dual case (i.e. negative lateral functors) holds in exactly the same way.

Lemma 3.11 motivates the following claim:

$$
[\mathbb{C}, \mathbb{D}] \sim \bigcup_{n \in N} \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right] \cup \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{n} \mathbb{D}\right]
$$

[^57]Remark. As $T$ is full and faithful, there is an equivalence (in ML) between $\operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, T^{m} \mathbb{D}\right]$ (where $n>m$ ) and $\operatorname{cat}(\mathcal{N})\left[T^{n-m} \mathbb{C}, \mathbb{D}\right]$. The dual case (i.e. where $n<m$ ) is equally clear. Thus, we are justified in excluding those (internal) functor categories $\operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, T^{m} \mathbb{D}\right]$ (where $n, m \neq 0$ ) from our proposed external functor category.

## Limitations on Inhomogeneity

As $I$ is a genuine class function of ML, taking a union of iterated applications of $T$ to a given set/category is not problematic in the way we might expect, were we working within the confines of NF. But there remains an outstanding impediment to the equivalence:

$$
[\mathbb{C}, \mathbb{D}] \cong \bigcup_{n \in N} \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right] \cup \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{n} \mathbb{D}\right]
$$

Were the equivalence to hold, it would imply that no natural transformations can exist between external functors of differing degrees. In other words, the above equation depends on the statement:

$$
\text { Given } F, G \in[\mathbb{C}, \mathbb{D}] .\left(F ' n \text {-lateral' } \wedge \neg G{ }^{\prime} n \text {-lateral' }\right) \Longrightarrow[\mathbb{C}, \mathbb{D}](F, G)=\emptyset
$$

A natural transformation $\tau \in[\mathbb{C}, \mathbb{D}](F, G)$ is defined as a collection of morphisms:

$$
\tau=\left\{\tau_{c} \in D_{1} \mid c \in C_{0} \wedge \tau_{c}: F c \rightarrow G c\right\}
$$

satisfying the naturality condition that the following diagram commutes for any morphism $f$ in $\mathbb{C}$ :


Remark (The Definable Case). $F, G: \mathbb{C} \rightarrow \mathbb{D}$ are said to be definable by stratified formulae $\Phi$ and $\Psi$, respectively, when the following correspondence holds for $i \in\{0,1\}$ :

$$
\begin{aligned}
& x \mapsto F_{i}(x)=\left\{y \mid \Phi_{i}(y, x)\right\} \\
& x \mapsto G_{i}(x)=\left\{z \mid \Psi_{i}(z, x)\right\}
\end{aligned}
$$

Even if $F$ and $G$ are proper classes, the natural transformations between them must be collections of homogeneous functional relations. Thus any component $\tau_{c}$ of a natural transformation $\tau: F \rightarrow G$ is a functional sub-relation of :

$$
\left\{\langle y, z\rangle \mid \Phi_{0}(y, c) \wedge \Psi_{0}(z, c)\right\}
$$

But we seek to prove a more general result, with no reference to syntax or definability.

Lemma 3.12. A natural transformation is only definable between (external) functors with equivalent degrees of inhomogeneity.

Proof. The content of this statement is: given (external) functors $F, G: \mathbb{C} \rightarrow \mathbb{D}$ and a natural transformation $\tau: F \rightarrow G$, there exists some $n$ such that both $F$ and $G$ are $n$-lateral. Recall, $\tau$ is a collection of sets:

$$
\left\{\tau_{c}: F c \rightarrow G c \mid c \in C_{0}\right\} \text { such that: } \forall f \in C_{1} \cdot \tau_{c o d(f)} \circ F f=G f \circ \tau_{\operatorname{dom}(f)}
$$

Both $F$ and $G$ are lateral, thus for some $n$, both $F_{0}$ and $F_{1}$ are $n$-lateral functions

$$
\exists n \in N \cdot \operatorname{set}\left(\left\{\left\langle\iota^{n} c, F^{c} c\right\rangle \mid c \in C_{0}\right\}\right)
$$

But this implies the function indexing each (set) morphism $\tau_{c}$ in $\tau$ is itself $n$-lateral,

$$
\operatorname{set}\left(\left\{\left\langle\iota^{n!} c, \tau_{c}\right\rangle \mid c \in C_{0}\right\}\right)
$$

From which we obtain immediately (as each $\tau_{c}$ is a set),

$$
\operatorname{set}\left(\left\{\left\langle\iota^{n \iota} c, G^{6} c\right\rangle \mid c \in C_{0}\right\}\right)
$$

Thus, $G$ must also be $n$-lateral. The "dual" case, where $F$ is $-n$-lateral, follows in exactly the same way.

Corollary 3.13. For any pair of $\mathcal{N}$-internal categories, $\mathbb{C}$ and $\mathbb{D}$, there is an equivalence of categories:

$$
[\mathbb{C}, \mathbb{D}] \cong \bigcup_{n \in N} \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right] \cup \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{n} \mathbb{D}\right]
$$

Proof. As natural transformations exist only between (external) functors of equivalent degrees of inhomogeneity (i.e. $\exists n . F, G$ are $n$-lateral), $[\mathbb{C}, \mathbb{D}]$ consists of countably many path-connected components, denoted $[\mathbb{C}, \mathbb{D}]_{n}$, for each which corresponds to the $n$-lateral functors between $\mathbb{C}$ and $\mathbb{D}$, for some $n \in \mathbb{Z}$. Lemmas 3.11 and 3.12 imply the following equivalences:

$$
\begin{array}{lr}
{[\mathbb{C}, \mathbb{D}]_{n} \cong \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right]} & (n>0) \\
{[\mathbb{C}, \mathbb{D}]_{n} \cong \operatorname{cat}(\mathcal{N})[\mathbb{C}, \mathbb{D}]} & (m<0) \\
{[\mathbb{C}, \mathbb{D}]_{m} \cong \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{-m} \mathbb{D}\right]} & (m<0)
\end{array}
$$

Thus, we have a partition of $[\mathbb{C}, \mathbb{D}]$ into subcategories, exactly as described by the category:

$$
\bigcup_{n \in N} \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right] \cup \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{n} \mathbb{D}\right]
$$

It is important to note that the equivalence of categories,

$$
[\mathbb{C}, \mathbb{D}]_{n} \cong \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right]
$$

corresponds to precomoposition with isomorphism of categories: $\tilde{I}_{n}: \mathbb{C} \rightarrow T^{n} \mathbb{C}$, in ML. Indeed, despite the apparent oddity of our (external) functor categories, we obtain an important general property.

Proposition 3.14. Given categories $\mathbb{C}, \mathbb{D}$ and a n-lateral functor $F: \mathbb{E} \rightarrow \mathbb{C}$, we obtain a $n$-lateral functor:

$$
[\mathbb{C}, F]:[\mathbb{C},-] \rightarrow[\mathbb{E},-]:[\mathbb{C}, \mathbb{D}] \mapsto[\mathbb{C}, \mathbb{E}]
$$

induced by pre-composition: $[\mathbb{C}, F]:(G: \mathbb{C} \rightarrow \mathbb{D}) \mapsto(G \circ F: \mathbb{E} \rightarrow \mathbb{D})$.

Proof. Recall that the composite of an $n$-lateral function, $f$, with an $m$-lateral function, $g$, is an $m+n$-lateral function $g \circ f$. Of course, this also holds true of lateral functors. As $F: \mathbb{E} \rightarrow \mathbb{C}$ is $n$-lateral, there is a corresponding internal functor, $\tilde{F}: T^{n} \mathbb{E} \rightarrow \mathbb{C}$. Given any lateral functor $G: \mathbb{C} \rightarrow \mathbb{D}$, we need to consider two cases.
$(m>0)$ If $G$ is $m$-lateral, there is a corresponding internal functor $\tilde{G}: T^{m} \mathbb{C} \rightarrow \mathbb{D}$. $G \circ F: \mathbb{E} \rightarrow \mathbb{D}$ is then the $n+m$-lateral functor corresponding to the composite of internal functors:

$$
\widetilde{G \circ F}=\tilde{G} \circ T^{m} \tilde{F}: T^{n+m} \mathbb{E} \rightarrow \mathbb{D}
$$

Thus, for $m>0,[\mathbb{C}, F]$ can be defined component-wise by the $n$-lateral functor between (internal) functor categories:

$$
[\mathbb{C}, F]_{m}: \operatorname{cat}(\mathcal{N})\left[T^{m} \mathbb{C}, \mathbb{D}\right] \rightarrow \operatorname{cat}(\mathcal{N})\left[T^{n+m} \mathbb{E}, \mathbb{D}\right]
$$

$(m<0)$ If $H$ is $m$-lateral, there is a corresponding internal functor $\tilde{H}: \mathbb{C} \rightarrow T^{-m} \mathbb{D}$. We claim $H \circ F: \mathbb{E} \rightarrow \mathbb{D}$ is then an $n+m$-lateral functor, but the obvious composite of internal functors is: $\tilde{H} \circ \tilde{F}: T^{n} \mathbb{E} \rightarrow T^{-m} \mathbb{D}$. By convention, therefore, we need to consider three cases:

$$
\begin{array}{ll}
\widetilde{H \circ F}=T^{m}(\tilde{H} \circ \tilde{F}): T^{(n+m)} \mathbb{E} \rightarrow \mathbb{D} & (m+n>0) \\
\widetilde{H \circ F}=T^{m}(\tilde{H} \circ \tilde{F}): \mathbb{E} \rightarrow \mathbb{D} & (m+n=0) \\
\widetilde{H \circ F}=T^{-n}(\tilde{H} \circ \tilde{F}): \mathbb{E} \rightarrow T^{-(n+m)} \mathbb{D} & \\
(m+n<0)
\end{array}
$$

Thus, for $m<0,[\mathbb{C}, F]$ can be defined component-wise by the $n$-lateral functor between (internal) functor categories:

$$
\begin{array}{ll}
{[\mathbb{C}, F]_{m}: \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{-m} \mathbb{D}\right] \rightarrow \operatorname{cat}(\mathcal{N})\left[T^{n+m} \mathbb{E}, \mathbb{D}\right]} & (m+n>0) \\
{[\mathbb{C}, F]_{m}: \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{-m} \mathbb{D}\right] \rightarrow \operatorname{cat}(\mathcal{N})[\mathbb{E}, \mathbb{D}]} & (m+n=0) \\
{[\mathbb{C}, F]_{m}: \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{-m} \mathbb{D}\right] \rightarrow \operatorname{cat}(\mathcal{N})\left[\mathbb{E}, T^{-(n+m)} \mathbb{D}\right]} & (m+n<0)
\end{array}
$$

Thus, we obtain an $n$-lateral functor, induced by precomposition with $F$, defined by the action of $[\mathbb{C}, F]_{m}$ on each compoent of

$$
[\mathbb{C}, \mathbb{D}]=\bigcup_{n \in N} \operatorname{cat}(\mathcal{N})\left[T^{n} \mathbb{C}, \mathbb{D}\right] \cup \operatorname{cat}(\mathcal{N})\left[\mathbb{C}, T^{n} \mathbb{D}\right]
$$

Remark (Other Functors in ML). One could easily form the category of ML classes (call it $\mathcal{M})$ and show that $\operatorname{cat}(\mathcal{N}) \subset \operatorname{cat}(\mathcal{M})$. But even if we expanded the functor categories
of $\operatorname{cat}(\mathcal{N})$ to be those consisting of all lateral functors (i.e. those of the form $[\mathbb{C}, \mathbb{D}]$, as defined above), we would not obtain a full subcategory of $\operatorname{cat}(\mathcal{M})$. In other words, by admitting only lateral functors, we are being more restrictive than if we were to simply take the generic definition of an $\mathcal{M}$-internal functor.

Nevertheless, by making this restriction we obtain some important "coherence" properties:

1. Given a pair of $\mathcal{N}$-internal categories $\mathbb{C}$ and $\mathbb{D}$, their external functor category $[\mathbb{C}, \mathbb{D}]$ is equivalent to a $\mathcal{M}$-internal category.
2. While lateral functors are semantic objects (i.e. there is no reference to syntax/definability), we can employ them, sensibly, in (stratified) formulae.
3. As the language of category theory is homogeneous, any structure we wish to define diagramatically will result in a lateral functor.

Remark (Returning to NF). From the above results, one might be tempted to think that rather than develop category theory in $\mathcal{N}$, we should develop it in the category of ML classes. Ultimately, however, the primary category theoretic relevance of ML is simply to prove that external functors/categories exist as genuine objects of our metatheory ${ }^{22}$

For example, working generally in $\mathcal{M}$, we would lose a number of closure properties present in $\mathcal{N}$. Unlike the category of NF sets, $\operatorname{cat}(\mathcal{M}) \notin \operatorname{cat}(\mathcal{M})$, and internal functor categories (i.e. $\operatorname{cat}(\mathcal{M})[\mathbb{C}, \mathbb{D}])$ would not exist as objects of the ambient category $\mathcal{M}$. Furthermore, even the admission of $\tilde{I}_{n}$ as an internal (isomorphic) functor would fail to produce adjunctions, in place of the symmetric lifts we obtain in $\mathcal{N}$. While lateral functors allow us to define an (internal) isomorphism between $\mathcal{N}$ and $T^{n} \mathcal{N}$ but, despite the fact that $T^{n} \mathcal{N}$ is a proper subcategory of $\mathcal{N}$, this does not allow us to assert that $\tilde{I}_{n}$ extends to an isomorphic endofunctor on $\mathcal{N}$ - the inverse, $\tilde{I_{-n}}$, is simply not defined

[^58]on all objects of $\mathcal{N}{ }^{23}$

Thus, the remainder of this thesis will work solely in the context of NF, seeking to develop the category theory in and of $\mathcal{N}$.

### 3.1.3 The T Functor

Any category of stratified sets has some residue of "type-shifting," which we refer to as the $T$-functor. The instantiation of $T$ in $\mathcal{N}$ and $\mathcal{K}$ is straightforward. But it is also important to think about the appropriate general definition, both by examining candidates for $T$-like functors from other areas of category theory (e.g. the Yoneda embedding) and by examining the implications (for $\mathcal{N}$ and in general) of $T$ possessing particular properties.

## Intuition and Basic Properties for T

For a given formula, expressible in the language of set theory, we can determine its stratified analogue by making the minimal adjustments necessary to obtain a stratified formula. For example, while we cannot represent the $\in$ relation as $\{\langle x, y\rangle \mid x \in y\} \subset$ $V \times V$, we can represent the relation $\{\langle\{x\}, y\rangle \mid x \in y\}$. The fundamental challenge in NF (or any stratified theory) is not type-raising, but type-lowering. The obvious typelowering operation, $\bigcup$, is not injective ${ }^{24}$ The type-raising operation $\iota: x \mapsto\{x\}$ is, as witnessed by the fact that $\bigcup$ is a left inverse. The more category theoretic interpretation of $\bigcup \circ \iota=i d$ is: $T$ is a full and faithful embedding.

TST has a clear notion of rank: $x \in T_{n}$ has rank $n$. We can use this information and always type-raise a variable $x_{i}$, by applying ' $\iota$ ' $n$ levels down (denoted $\left.{ }^{25} j^{n}(\iota)^{\prime} x\right)$. By

[^59]applying $\iota$ sufficiently far down, one preserves all categorical structure.

NF does not furnish a well-defined notion of rank. So there is no canonical choice of how many levels down one should apply $\iota$, to preserve necessary structure. This does not impact the action of $T$ on morphisms, as it is defined by a distinct operation on $|\mathcal{N}|$ and $\operatorname{Mor}(\mathcal{N}): T_{0} \sim \iota^{\prime \prime}$ and $T_{1} \sim R U S C$. In general, $T$ preserves structures at least up to isomorphism. Here one uses the canonical $N$-indexed isomorphism class of $T$-functors: $\left\{j^{n}(\iota) \mid n \in N\right\}$ and $\left\{j^{n}(R U S C) \mid n \in N\right\}$.

We verify the basic properties of $T$, which hold in $\mathcal{N}$ and $\mathcal{K}$ :

Lemma 3.15. $T$ is a full and faithful functor, but not essentially surjective.

Proof. By definition, $T$ is a full and faithful embedding of $\mathcal{N}$ into itself. Essential surjectivity fails by Cantor's Theorem: $|T V|<|V|$.

Lemma 3.16. $T$ creates finite limits.

Proof. We know that $T$ preserves finite limits. As $T$ is full and faithful, it reflects finite limits (and finite colimits, where they exist) ${ }^{26}$

Remark. The action of $T$ on morphisms is defined as $R U S C$. In the case of function spaces (treated as objects) $T$ is equipped with a natural isomorphism:

$$
e_{X, Y}:(T X \Rightarrow T Y) \cong T(X \Rightarrow Y)
$$

For the instantiation of $T$ in $\mathcal{N}$, the existence of $e_{X, Y}$ is pure formality. But as we consider a broader variety of potential $T$-functors, it is worth keeping track of the need for certain "coherence" conditions (See Definition 3.18).

[^60]
## Alternative T Functors

## $T$ and $\mathcal{N}(1,-)$

Classically, given a locally small category $\mathcal{C}$ with a terminal object, we can embed it into Set, by taking the function $\mathcal{C}(1,-): \mathcal{C} \rightarrow$ Set. As Set is itself locally small, we can consider the Freyd cover of $\operatorname{Set}, X \in \operatorname{Set} \mapsto \operatorname{Set}(1, X)$ [28]. In classical set theory, $\operatorname{Set}(1, X) \cong X$.

In $\mathrm{NF}, \mathcal{N}(1,-): \mathcal{N} \rightarrow \mathcal{N}$ is a valid (external) functor, but we obtain an isomorphism ${ }^{[27}$

$$
T X \cong \mathcal{N}(1, X)
$$

The bijection extends to a natural isomorphism between $\mathcal{N}(1,-)$ and $T$. This motivates a somewhat counterintuitive corollary to Lemma 3.15. Despite $T V$ being strictly smaller than $|V|$, the equivalence of a given pair of parallel morphisms is preserved (and reflected) under $T$.

Lemma 3.17. $\mathcal{N}$ is well-pointed.

Remark (NF Curry-Howard). The classical cartesian closure adjunction, $A \times(-) \dashv(-)^{A}$, is equivalent to the rule of deduction:

$$
(A \wedge B \vdash C) \Longleftrightarrow(B \vdash A \Rightarrow C)
$$

In $\mathcal{N}$, we have the deduction rule:

$$
(A \wedge B \vdash C) \Longleftrightarrow((1 \Rightarrow B) \vdash(A \Rightarrow C))
$$

Thus, we have a logic where ' $A$ and $B$ entail $C$ ' if and only if ' $B$ is provable entails $A$ implies $C$.' The obvious isomorphism, $(1 \Rightarrow B) \cong T B$ permits an interpretation of $T B$ as ' $B$ is provable,' arising from the type-theoretic interpretation ' $B$ is inhabited.' Morphisms $1 \rightarrow B$ correspond to terms or, equivalently, proofs of $B$.

[^61]
## Yoneda Embedding

The relationship between the Yoneda Embedding and $T$ was discussed in Section 2.2.1. This follows the relationship between $\{\cdot\}: 1 \Rightarrow P$ and $\mathcal{Y}$, but is slightly more nuanced. Viewing presheaves as a generalization of powersets, both $\{\cdot\}$ and $\mathcal{Y}$ arise as units to a "powerset" monad. In the classical case (i.e. category theory over Set) this relationship is not entirely precise - due to size restrictions, the presheaf functor corresponds only to a relative (pseudo)monad.

In $\mathcal{N}$, on the other hand, $P$ and $\widehat{(-)}$ are genuine instances of the same categorical structure. $P$ and $\widehat{(-)}$ are both restricted due to inhomogeneity, and neither is restricted due to size (i.e. both form relative (pseudo)monads). Thus, $\mathcal{Y}$ can be viewed as a "higher" (in the category theoretic sense - where we view sets as discrete categories) form of $T$. Both arise as units to a relative (powerset) monad, the only distinction being that $T$ is $\mathcal{Y}$ restricted to discrete categories (i.e. sets).

## An Isomorphism Class of T Functors

The original motivation for $T$ was the cardinal arithmetic of NF:

$$
\alpha \mapsto T \alpha=\{x \mid \exists y \in \alpha \cdot x \cong \iota " y\}
$$

Implementation of $T$ as a functor requires a selection (the standard one being $\iota^{\prime \prime}$ ) among an isomorphism class of possible $T$-functors. Not only do we have an isomorphism $j^{n}(\iota) \cong j^{n+1}(\iota)$, we have a unique natural isomorphism between functors. A previously noted advantage of this is that any $n$-stratifiable property is preserved under the embedding $j^{n}(\iota)$.

Much to our surprise, $j^{n}(\iota) \cong j^{n+1}(\iota)$ does not appear to be purely set theoretic. In the following section, we see that it is implied by the fact that (syntactic) coequalizers in $\mathcal{N}$ form a symmetric lift ${ }^{28}$

[^62]
### 3.1.4 A More General Category

It would be misleading to say that there is a canonical category, from which the general notion of a topos is abstracted [23]. But it is hard to ignore the role of Set as a "motivating" example ${ }^{29}$ Classically, Set is an elementary topos. In the other direction, the internal language of a topos allows one to express the (local) $\in$-relation. Therefore, while we should be cautious not to think of a topos as "depending" on a chosen base category of sets, given that we are considering an alternative category of sets, we should consider the possibility of a corresponding alternative topos.

A sensible "NF-topos" would include a $T$-functor and the ability to express stratified formulae in its internal language. We do not require the existence of a universal set - this would exclude many models of KF - but there should be a consistent means of admitting such an object to an "NF-topos," without forming a paradoxical structure ${ }^{30}$ In an NF-topos there must be an interdependency between the existence of a universe object and certain properties of $T$ (Theorem 3.34) ${ }^{31}$

An NF-topos, which we refer to as an $S P E$, generalizes the definition of both an elementary topos (recall, Mac is an extension of $K F$ ) and $\mathcal{N}$.

Definition 3.18. [14] An SPE is a category $\mathcal{C}$ such that:

1. $\mathcal{C}$ is a regular category with finite coproducts and a subobject classifier
2. There is a full and faithful embedding $T: \mathcal{C} \rightarrow \mathcal{C}$ which creates finite limits
3. There is a bifunctor $\Rightarrow: \mathcal{C}^{o p} \times \mathcal{C} \rightarrow \mathcal{C}$ with the following coherence isomorphisms:
(a) A natural isomorphism $i_{A}: T A \cong 1 \rightarrow A$.
(b) An extranatural transformation $\alpha_{A}: 1 \rightarrow A \Rightarrow A$.

[^63](c) A transformation:
$$
\beta_{A, B, X}: T A \rightarrow T B \Rightarrow(X \rightarrow A) \rightarrow(X \rightarrow B)
$$ natural in $A$ and $B$, extranatural in $X$.
(d) A natural isomorphism $e_{A, B}:(T A \rightarrow T B) \cong T(A \rightarrow B)$.
4. For all $f: A \rightarrow B$, there is a functor $\tilde{\Pi}_{f}: \mathcal{C} / A \rightarrow \mathcal{C} / T B$, for which $f^{*} T_{B} \dashv \tilde{\Pi}_{f}{ }^{32}$
5. The functor $\operatorname{Sub}(T A \times-)$ is representable for any $A^{33}$

### 3.2 Basic Categorical Properties

This section proves some general properties of categories of stratified sets. Whenever it is possible, we prove results in KF . But $\mathcal{N}$ remains our primary category of interest. The category of NF sets is shown to be a regular category with a natural numbers object.

Remark (Interpreting Concrete Finiteness). While every concrete finite set is strongly cantorian, proving the statement "all finite sets are strongly cantorian" requires the Axiom of Counting. Similarly, when we make a statement like " $\mathcal{N}$ has finite limits," we mean "for any concretely finite diagram, we can form its limit in $\mathcal{N}$."

The introduction of Quine Sequences (Definition 5.57) allows us to go some way towards dropping the concreteness condition. But there is no clear disadvantage to interpreting "finite completeness" as "given a diagram of size ' $n$,' it has a limit in $\mathcal{N}$."

### 3.2.1 KF

For most of this section we work in $\mathrm{KF}(\mathrm{I})$ but, for the following proposition, we consider the alternative extension, $\mathrm{KF}+\mathbf{I O}$. As a consequence, stratified $\Delta_{0}$-separation becomes less restrictive. Nevertheless, restricted quantification forces one to work locally (i.e. all sets assumed in the hypothesis occur within a pre-existing set).

[^64]Proposition 3.19. In $K F+I O$, finite products and coproducts exist (locally).

Proof. Given two sets, $X$ and $Y$, both contained in some larger set $Z$, we can form a binary product $X \times Y$, with projection morphisms whose graphs are sets. As we do not assume infinity, we implement Kuratowski ordered pairs. To achieve projection functions, definable as sets, we must invoke IO two times. Two iterations of IO give a bijective map $f$, defined on $Z$, such that $\forall z \in Z \exists b \cdot f(z)=\{\{b\}\}$. The product $X \times Y$ will be:

$$
\left\{\left\langle\cup^{2}(f(x)), \cup^{2}(f(y))\right\rangle \mid x \in X \wedge y \in Y\right\}
$$

As $\cup^{2}(f(x))$ is two types below $x$ and Kuratowski pairs are two types above their components, the graphs of projection functions are sets. This can be extended to $n$-indexed products by $n$ applications of IO. Coproducts are defined similarly.

Lemma 3.20. KF has equalizers.

Proof. Given a set $X$ and two morphisms $f, g: X \rightarrow Y,\{x \mid f(x)=g(x)\}$ receives the same type as $X$ in any valid stratification.

Corollary 3.21. $K F+I O$ has finite limits.

The existence of a subobject classifier is also unproblematic, and does not even require IO.

Remark (Uniqueness over Bijectivity). For any set $X$ in KF (or NF) and any injection $m: A \hookrightarrow X$, there is a canonical morphism $\chi_{m}: X \rightarrow \Omega$, sending an element $x \in X$ to $T$ precisely when it is in the image of $m$. In other words, fitting into the pullback diagram below:


In a stratified theory, the existence of such a universal property does not always imply a bijection between hom-sets. In this case, however, the typing is well-behaved and we can prove a bijection between $\left\{i d_{x} \mid x \subset X\right\}$ and $\{\chi \mid \chi: X \rightarrow \Omega\}$.

Many other examples of universal properties do, indeed, define bijections between homsets, but we do not state this as a general requirement. A number of relative adjunctions in $\mathcal{N}$ satisfy a universal property, but are not part of an internal bijection of hom-sets. In this sense, we emphasize uniqueness over bijectivity.

## Coequalizers and Partitions in KF(I)

Implementation of type-level pairs in $\mathrm{KF}(\mathrm{I})$ permits the formation of finite limits and coproducts in the expected way - the corresponding cartesian products and disjoint unions have the desired universal properties. However, restriction to stratified $\Delta_{0}$-separation is problematic for the formation of colimits.

The complication arises from the implicit type-raising in the formation of equivalence classes. The classical set theoretic definition of colimits makes implicit use of quotient sets. In a stratified theory such as $\mathrm{KF}(\mathrm{I})$, given some equivalence relation $R \rightharpoondown X \times X$ on a set $X$, the set of equivalence classes under $R,\left\{[x]_{R} \mid x \in X\right\}$, is one type higher than $X$ in any valid stratification.

It is helpful to differentiate between semantic (universal) properties and the canonical synactic implementation of the property in Set. Formally, what we refer to as "semantic" coincides with the standard definition.

Definition 3.22. Given $f, g: X \rightarrow Y$, a semantic coequalizer, requires a morphism $c: Y \rightarrow Z$, satisfying the universal property: given any $h: Y \rightarrow W$ such that $h \circ f=h \circ g$, there exists a unique morphism $\bar{h}: Z \rightarrow W$ such that $\bar{h} \circ c=h$.

To obtain such a map in Set, $Z$ is defined as the quotient of $Y$, under the $\subseteq$-least equivalence relation extending:

$$
\sim_{f, g}=\left\{\left\langle y, y^{\prime}\right\rangle \mid \exists x \in X . y=f(x) \wedge y^{\prime}=g(x)\right\}
$$

$c$ is then just the map sending each element to its equivalence class. We refer to this as the syntactic coequalizer ${ }^{34}$

[^65]Definition 3.23. Given a pair of parallel maps $f, g: X \rightarrow Y$, we refer to the map:

$$
c_{f, g}: Y \rightarrow Y / \sim_{f, g}=\operatorname{coeq}(f, g) ; y \mapsto[y]_{\sim}
$$

as a strongly syntactic coequalizer ${ }^{35}$

In the instance that a (semantic) coequalizing map exists, but is not necessarily $c_{f, g}$, and its target is isomorphic to $\operatorname{coeq}(f, g)$, we will refer to the map as merely a syntactic coequalizer. Here we seek to draw a distinction similar to that between cantorian and strongly cantorian.

In $\operatorname{KF}(\mathrm{I})$, the set $\operatorname{coeq}(f, g)$ exists, but the graph of $c_{f, g}$ is unstratified. In the extension $\mathrm{KF}+\mathbf{I O}$, one can "correct" the typing mismatch, as displayed in the diagram below, where the coequalizing map arises as the "composite" of two inhomogeneous operations.


One might hope for a weaker extension of $\mathrm{KF}(\mathrm{I})$ in which the (syntactic) coequalizer diagram could still be completed, generally. While the quotient object $Z$ is a subset of $P Y$ (so one type higher than $Y$ ), it is a subset that satisfies a particular property: the elements of $Z$ are pairwise disjoint. Therefore, IO need only hold for pairwise disjoint families of sets - precisely the axiom CE, defined previously. In fact, the extension of $\mathrm{KF}(\mathrm{I})$ to $\mathrm{KF}(\mathrm{I})+\mathbf{C E}$ is exactly the correct strength.

Lemma 3.24. For any set $Y$ and any partition $Z$ of $Y$, we can prove the existence of two morphisms $f, g: X \rightarrow Y$ such that the partition $Z$ is the set of equivalence classes of $Y$, corresponding to the $\subseteq$-least equivalence relation extending the relation:

$$
\sim_{f, g}=\left\{\left\langle y, y^{\prime}\right\rangle \mid \exists x \in X . y=f(x) \wedge y^{\prime}=g(x)\right\}
$$

[^66]Proof. Form the (equivalence) relation on $Y$ that induces the partition $Z$ :

$$
\left\{\left\langle y, y^{\prime}\right\rangle \mid \exists z \in Z . y \in z \wedge y^{\prime} \in z\right\}
$$

This has the same type as $Y$, in any stratification. So $f$ and $g$ are just the two projection functions.

Theorem 3.25. In $K F(I)$, the following are equivalent:
(i) Every (syntactic) coequalizer diagram can be completed.
(ii) Every set of pairwise disjoint sets is the same size as a set of singletons.

Proof. (ii) $\Longrightarrow$ (i): Given two morphisms $f, g: X \rightarrow Y$, the $\subseteq$-least equivalence relation defined above:

$$
\left\{\left\langle y, y^{\prime}\right\rangle \mid \exists x \in X . y=f(x) \wedge y^{\prime}=g(x)\right\}
$$

defines a partition $Z$ of $Y$. By (ii), such a partition $Z$ is the same size as a set of singletons, $\iota$ " $C$. So we have a definable bijection $h: Z \rightarrow \iota " C$. Thus, we can form the coequalizing morphism:

$$
c: Y \rightarrow C ; y \mapsto \cup h([y])
$$

(i) $\Longrightarrow$ (ii): Let $\Pi$ be a set of pairwise disjoint sets and form $Y=\cup \Pi$. Lemma 3.24 implies we can form $f, g: X \rightarrow Y$, such that $\Pi$ is the induced partition. What we desire is, effectively, a "choice set" $C$, with exactly one element corresponding to each equivalence class on $Y$, induced by $\Pi$. $C$ is then the appropriate "type" to form a map sending an element $y \in Y$ to the $c \in C$, which corresponds to its equivalence class. ${ }^{36}$ (i) implies that $f$ and $g$ fit into a coequalizer diagram, with a coequalizing morphism given by $c: Y \rightarrow Z$, where $Z$ has the property required of $C$. The inverse of $c$ is seen to restrict to a bijection from $\iota$ " $Z \rightarrow \Pi$. Thus, (ii) is proven.

The above result forms a connection between the existence of coequalizers and the axiom of choice in a stratified theory. Choice for pairwise disjoint families of sets implies $\mathbf{C E}$

[^67]and, by corollary, the existence of coequalizers. It is known that choice fails in NF, but we can obtain a corollary for NFU + Choice:

Corollary 3.26. The category of sets defined in NFU + Choice has coequalizers.
Remark (Choice and Stratified Theories). Choice sets (and their corresponding maps), where they exist, play an important role in stratified set theories - they allow one to "type-lower" families of sets that satisfy the conditions placed on AC. Axioms such as IO and CE do not imply the existence choice maps per se, but they do imply the existence of what we refer to as selection maps - effectively, the ability to state that a given set is in bijection with a set that has trivial choice maps (i.e. a set containing only singleton sets).

IO turns out to be very strong - hence, refutable in NF. On the other hand, there is no obvious reason to doubt the consistency of NF $+\mathbf{C E}$.

### 3.2.2 NF

The category of NF sets inherits what has been proven above for $\mathrm{KF}(\mathrm{I}){ }^{37}$
Proposition 3.27 (Summary). $\mathcal{N}$ has finite limits, a subobject classifier and finite coproducts.

We can prove the existence of certain coequalizers.
Lemma 3.28. $\mathcal{N}$ has coequalizers of kernel pairs

Proof. The kernel pair of $f$ corresponds to the projection morphisms for the equivalence relation:

$$
\sim_{f}=\left\{\left\langle y, y^{\prime}\right\rangle \mid f(y)=f\left(y^{\prime}\right)\right\}
$$

Unlike arbitrary partitions, however, the fibres of a given function are in canonical bijection with a set of singletons, as witnessed by $f^{-1}$. The coequalizing morphism is induced by $f$, which acts as the canonical (trivial) selection function.

[^68]The same argument shows that every epimorphism in $\mathcal{N}$ is regular.
Proposition 3.29. $\mathcal{N}$ is a regular category.
$\mathrm{NF}+\mathbf{C E}$, as an extension of $\operatorname{KF}(\mathrm{I})+\mathbf{C E}$, has coequalizers and, as a corollary to the above results, is finitely cocomplete.

Corollary 3.30. $\mathcal{N}_{C E}$ is a regular category with finite colimits.

The internal category theory (specifically, the internal presheaves) of $\mathcal{N}_{C E}$ also turns out to have stronger representative power than $\operatorname{cat}(\mathcal{N})$.

Absent CE, it is not obvious that $\mathcal{N}$ does not have coequalizers for arbitrary pairs. We can, however, prove that no consistent extension of NF has syntactic coequalizers in the strict sense (i.e. there exist syntactic coequalizer diagrams that cannot be completed by a morphism with the action: $\left.\lambda_{x}(\imath y)(x \in y)\right)$.

Theorem 3.31. For any consistent extension of $N F$, the category $\mathcal{N}$ does not have strongly syntactic coequalizers. In other words, for arbitrary parallel pairs $f, g: X \rightarrow Y$.

$$
c_{f, g}: Y \rightarrow Y / \sim_{f, g}=\operatorname{coeq}(f, g) ; y \mapsto[y]_{\sim_{f, g}}
$$

Proof. Lemmas 3.24 and 3.32 .
Lemma 3.32. NF refutes that every pair of parallel projections, constructed as in Lemma 3.24, can be coequalised by the quotient set equipped with the morphism $\lambda_{x}(\imath y)(x \in$ $y)$.

Proof. Let $B$ be the set of all wellorderings and let $A$ be the relation:

$$
\{\langle x, y\rangle \mid x, y \in B . x \text { is order-isomorphic to } y\}
$$

The quotient of $A$ 's projections is just $N O$, the ordinal numbers. Define a function $c$ :

$$
c \equiv \lambda_{x}(\imath y)(x \in y): B \rightarrow N O
$$

If the graph of $c$ is a set, then so is $j^{2}(c): N O \rightarrow \iota$ " $N O$. But each ordinal would then be the sole value of its members under $c$, meaning $j^{2}(c)$ is simply the singleton function. Such a singleton function would allow one to prove the Burali-Forti paradox.

Corollary 3.33. If $A$ is a family of non-empty, disjoint sets, then the existence of a membership morphism $\bigcup A \rightarrow A$ implies that $A$ is strongly cantorian.

## In $\mathcal{N}, \mathrm{T}$ is Neither an Adjoint nor a Monad

In [14], we sought to determine the appropriate general definition of a category of stratified sets, an SPE (Definition 3.18). In general, the properties of a given $T$-functor depend on (or are restricted by) the properties of its base category (the SPE). While the basic definition of $T$ is the same for $\mathcal{K}$ and $\mathcal{N}$, the stronger theory places restrictions on what properties $T$ can possess as a functor.

Theorem 3.34. In $\mathcal{N}, T$ is not part of an adjunction, nor is it a monad or co-monad.

Proof. (1) If $T$ were a right adjoint, there would be an injection from $\mathcal{N}(V, T 2)$ into $\mathcal{N}\left(T X_{V}, T 2\right)$, where $X_{V}$ is the object component of the $T$-universal arrow for $V$. However, $V \Rightarrow T 2 \cong P V$, as 2 is strongly cantorian. Thus, $|V \Rightarrow T 2| \cong|V|$. However, for any object $Y$ in $\mathcal{N},|T Y \Rightarrow T 2| \leq|T V|$. Thus we obtain a contradiction.
(2) If $T$ were a left adjoint, we would be able to prove the bijection between hom-sets:

$$
\mathcal{N}(T C, A \Rightarrow B) \cong \mathcal{N}(A \times C, B)
$$

If $T$ has a right adjoint $G$, we could prove:

$$
\mathcal{N}(T C, A \Rightarrow B) \cong \mathcal{N}(C, G(A \Rightarrow B))
$$

But in this case:

$$
\mathcal{N}(A \times C, B) \cong \mathcal{N}(C, G(A \Rightarrow B))
$$

would give a right adjoint to $A \times-$, contradicting the result that $\mathcal{N}$ is not cartesian closed.
(3) If $T$ were of the form $G F$ for $F \dashv G, F$ would be an embedding, as we have shown that $T$ is an embedding (i.e. a monomorphism in the category of functors). Thus, $F$ would be a faithful functor. Any adjunction with a faithful left adjoint has a monic unit
$\eta$. The component morphism $\eta_{V}: V \rightarrow G F V=T V$ would therefore be monic. As $|T V|<|V|$, this is a contradiction.
(4) Finally, if $T$ were of the form $F G$ for $F \dashv G, G$ would be faithful, so the co-unit $\varepsilon$ would be epimorphic. This would imply a surjective component morphism:

$$
\varepsilon_{V}: T V=F G V \rightarrow V
$$

again contradicting the result $|T V|<|V|$.

The following result corresponds to the relationship between $T$ and modified-cartesian closure (Definition 3.52) in an arbitrary SPE:

$$
\operatorname{hom}(T C, A \Rightarrow B) \cong \operatorname{hom}(A \times C, B)
$$

Corollary 3.35. Given an arbitrary SPE $\mathcal{C}$, the $T$-functor has a right adjoint if and only if $\mathcal{C}$ is a topos.

For the respective extensions of KF, NF and Mac, the relationship between $T$ and cartesian closure can be described by the diagram below, where $\mathcal{E}$ is an elementary topos:


## (Strongly Syntactic) Coequalizers as Symmetric Lifts

As previously observed, parallel pairs of morphisms between sets in the image of $T$ have (canonical) coequalizers. Equivalently, what we have referred to as the syntactic coequalizing functor coeq : $\mathcal{N} \rightrightarrows \rightarrow \mathcal{N}$ forms the relative adjunction coeq ${ }_{T} \rightrightarrows \cdot \dashv \Delta$. For maps $f, g: X \rightarrow Y$ :

$$
\mathcal{N}(\operatorname{coeq}(f, g), Z) \cong \mathcal{N}^{\rightrightarrows} \cdot\left((T f, T g), \Delta_{Z}\right)
$$

The coequalizing map for a pair of maps $T f, T g$ is given by the component of the relative unit:

$$
\iota_{f, g}:(T f, T g) \rightarrow \Delta_{\operatorname{coeq}(f, g)}=\left\langle 1_{\text {coeq }(f, g)}, 1_{\text {coeq }(f, g)}\right\rangle
$$

The relative adjunction, coeq $T_{T} \nrightarrow \dashv \Delta$, can thus be interpreted as a syntactic witness to the existence of semantic coequalizers for a (proper) subcategory of the category of diagrams of shape $\cdot \rightrightarrows \cdot$

It is straightforward to prove the existence of a second relative adjunction in $\mathcal{N}$, stating that $\Delta$ is $T$-right adjoint to coeq. coeq $\dashv_{T} \Delta$ corresponds to the natural isomorphism:

$$
\mathcal{N} \rightrightarrows \cdot\left((f, g), \Delta_{Z}\right) \cong \mathcal{N}(\operatorname{coeq}(f, g), T Z)
$$

The relative co-unit is given by its component for each object $Z{ }^{38}$

$$
\varepsilon_{Z}: \operatorname{coeq}\left(i d_{Z}, i d_{Z}\right) \rightarrow T Z
$$

The relative adjoints describing coequalizers are "coherent" in the sense that they paste to form a symmetric lift.


Proposition 3.36. The relative unit and co-unit of the corresponding relative adjunctions coeq $T_{T \rightrightarrows-} \dashv \Delta$ and coeq $\dashv_{T} \Delta$ paste to form the isomorphism:

$$
\text { coeq } \circ T \rightrightarrows . \cong T \circ \text { coeq }
$$

Proof. Given any two parallel maps $f, g: X \rightarrow Y$,

$$
\varepsilon_{\operatorname{coeq}(f, g)} \cdot \operatorname{coeq}\left(\iota_{(f, g)}\right): \operatorname{coeq}(T f, T g) \rightarrow T(\operatorname{coeq}(f, g))
$$

is an isomorphism, corresponding to a particular bijection, well-known to those who study NF, $j^{n}(\iota)(X) \cong j^{n+1}(\iota)(X)$. Explicitly $\varepsilon_{\text {coeq }(f, g)} \cdot \operatorname{coeq}\left(\iota_{(f, g)}\right)$ is:

$$
\operatorname{coeq}(T f, T g) \cong \operatorname{coeq}\left(1_{\operatorname{coeq}(f, g)}, 1_{\operatorname{coeq}(f, g)}\right) \cong T(\operatorname{coeq}(f, g)) ;[\{y\}]_{\sim} \mapsto\left\{[y]_{\sim}\right\}
$$

[^69]Remark (Symmetric Lifts and the Isomorphism Class of T Operations in NF). Proposition 3.36 is straightforward to prove, but it makes an important, explicit connection between coequalizers in categories of stratified sets and the canonical bijection:

$$
j^{n}(\iota)(X) \cong j^{n+1}(\iota)(X)
$$

Example 3.37 (Extending the Analogy Beyond NF). The result that the Yoneda Extension forms a symmetric lift also makes use of the fact that iteration of the Yoneda Embedding yields a series of fixed points ${ }^{39}$ We obtain the chain of bijections, natural in $C$ :

$$
F(C) \cong \mathcal{Y}_{\mathcal{C}}(F)(C) \cong \mathcal{Y}_{\widehat{\mathcal{C}}}\left(\mathcal{Y}_{\mathcal{C}}(F)\right)(C) \cong \ldots
$$

This implies $\mathcal{Y}$ and $T$ appear to be related by more than being relative units of presheaf (in the degenerate case, for $T$ ) relative monads.

Given some model of Zermelo set theory, a model of KF can be obtained as a direct limit of a model of TST, obtained by iterating the powerset function on a base set $X$ (i.e. $\left\{X, P X, P^{2} X, \ldots\right\}$ ). Sets of iterated singletons (i.e. $\left.\iota^{n}(x)\right)$ become identified in the same sense that the iterated Yoneda embedding is identified. If we conceive of a foundation for category theory, such as a Grothendieck Universe, the chain of Yoneda embeddings of a functor $F$ between small categories is effectively a collapse of functors between increasingly larger classes to a single functor $(F)$ between sets (the smallest level of classes) [31. Given that Mac Lane set theory is equivalent to the KF + "all sets are strongly cantorian" and corresponds to the internal set theory of a topos, it seems at least possible that KF is a sort of internal set theory for $C A T$, without a distinction between class sizes (i.e. with a form of stratified $\Delta_{0}$-separation for classes).

While this is an excessively casual conjecture, it highlights both the temptation and difficulty of generalizing the semantics of $\mathcal{N}$ to arbitrary categories.

[^70]
### 3.3 Universal Properties with Universal Sets

Links between the following three ideas have been discussed since the early days of category theory.

1. Universal Properties
2. The Axiom (Scheme) of Comprehension

## 3. The Axiom of Choice

Broadly speaking, in category theory, these ideas relate to the interaction between "syntax" and "semantics." Those who study NF have asked similar questions and made similar observations (e.g. consistent inclusion of a full scheme of stratified comprehension and failure of choice) 49].

There is a temptation (possibly correct) to say that semantics lead syntax. Taking this view to an extreme would be to consider the (syntactic) "set-building" method of separation/comprehension as nothing more than a choice function, among the isomorphism classes of objects in a category, satisfying a given universal property. A canonical (syntactic) choice then allows one to derive functorial universal properties from the more semantic conception of (saturated) anafunctors 34.

Still, comprehension is more than a method of "choosing" among existing isomorphism classes. In the words of Paul Taylor, there is a "generative" aspect of comprehension 67]. Even the most semantically minded individual would struggle to reduce the (categorical) study of Set to the study of "category theory with anafunctors + some choice principle among classes."

This section studies an alternative conception of universal properties in $\mathcal{N}$, permitted by unrestricted (stratified) comprehension. The implementations are similar to that of Frege ordinals - sets that are literal equivalence classes describing the order-type of their members.

## Implementing Universal Structures

Category theoretic properties are, in a naive sense, those which can be described diagrammatically. Informally, a property corresponds to the class of structures (within a given category) that satisfy it. For (locally) small categories, we can frequently express the category theoretic property in the language of our base set theory ${ }^{40}$ In this instance, a category theoretic property can be described formally as the class $\{\vec{x} \mid \Psi(\vec{x})\}$, where $\vec{x}$ ranges over maps and objects; and the binary relations expressing composition and equality in category theory are expressed in $\mathcal{L}_{\text {Set }}$.

Universal structures internalize a form of comprehension. For example, given the base $\xrightarrow{\lim } F$ of a universal cocone in $\mathcal{C}$, under a diagram $F$ (where $\Psi \sim$ 'is a cocone under $F$ '), one can form a bijection between classes:

$$
\{\vec{x} \mid \Psi(\vec{x})\} \cong\left\{f \in \operatorname{Mor}(\mathcal{C}) \mid d_{0}(f)=\underset{\longrightarrow}{\lim } F\right\}
$$

Universal structures are themselves defined only up to isomorphism, determining an equivalence class:

$$
\left\{f \in I \operatorname{so}(\mathcal{C}) \mid d_{0}(f)=\underset{\longrightarrow}{\lim } F\right\} \subset\{\vec{x} \mid \Psi(\vec{x})\}
$$

Classically (i.e. in $\mathrm{ZF}(\mathrm{C})$ ), to prove the existence of a category theoretic property described by $\Psi$, we seek to identify a representative object among this equivalence subclass of universal structures. In Set, we can obtain a representative universal $\Psi$-structure syntactically by determining an appropriate formula $\Phi$ and proving:

$$
\forall g \cdot[\Psi(g) \Longrightarrow \exists!h . h \circ\{\vec{z} \mid \Phi(\vec{z})\}=g]
$$

If we think about structure more abstractly, we can make $\Psi(\vec{x})$ more precise. Partition $\vec{x}$ into $\vec{z} \cup\{x\}$, where the structure $\Psi(\vec{z}, x)$ corresponds to an abstract "shape" and specific "images" correspond to the predicate $\Psi(\vec{a}, x)$, under substitution for variables in $\vec{z}$. In this context, the free variables of the syntactic representative universal $\Psi$-structure, $\Phi$, range over $\vec{z} \cup\left\{x^{\prime}\right\}$. Thus, we can think of the "shape" $\Psi$ as the abstract diagram $\mathcal{J}$, and $\vec{a}$ as an image of $\mathcal{J}$ in $S$ et (i.e. a functor). Thus, substitution of $\vec{a}$ for $\vec{z}$, in both

[^71]$\Psi$ and $\Phi$, induces a functorial universal $\Psi$-structure, where $\left\{x^{\prime} \mid \Phi\left(\vec{a}, x^{\prime}\right)\right\}$ is universal among members of the class $\{x \mid \Psi(\vec{a}, x)\}$.

Example 3.38 (Coequalizers in $S e t$ ). Given a pair of parallel morphisms $f, g$, we can define a series of formulas in $\mathcal{L}_{\text {Set }}$ :

$$
\begin{aligned}
& \operatorname{par}(f, g)=f, g \in \operatorname{Fun} \wedge d_{0}(f)=d_{0}(g) \wedge d_{1}(f)=d_{1}(g) \quad \text { (parallel maps) } \\
& \Psi(f, g, h)=h \in F u n \wedge h \circ f=h \circ g \wedge \operatorname{par}(f, g) \quad \text { (abstract (semantic) property) } \\
& \Psi_{f, g}=\{h \mid \Psi(f, g, h)\} \quad \text { (class of coequalizing maps at }(f, g) \text { ) } \\
& \sim_{f, g}=\left\{\left\langle z, z^{\prime}\right\rangle \mid \exists y \cdot f(y)=z \wedge g(y)=z^{\prime}\right\} \quad \text { ( } \subseteq \text {-least equivalence relation) } \\
& \Phi_{f, g}^{\prime}=\left\{w \mid \exists z \in d_{1}(f) \cdot[z]_{\sim_{f, g}}=w\right\} \quad \text { (universal object) } \\
& \Phi_{f, g}=\left\{\langle z, w\rangle \mid z \in d_{1}(f) \wedge z \in w \wedge w \in \Phi_{f, g}^{\prime}\right\} \quad \text { (universal map) } \\
& \forall k . \Psi(f, g, k) \Longrightarrow \exists!h . h \circ \Phi_{f, g}=k \\
& \text { (universal property) }
\end{aligned}
$$

We can state a weaker condition that proves the (potentially non-functorial) existence of universal $\Psi$-structures.

$$
\forall f, g \cdot[\operatorname{par}(f, g) \Longrightarrow \exists c .(\Psi(f, g, c) \wedge \forall k . \Psi(f, g, k) \Longrightarrow \exists!h . h \circ c=k)]
$$

In category theory, we typically prefer constructive proofs, as the internal language is constructive, by nature. But with unrestricted (stratified) comprehension, $\Psi_{f, g}$ determines a set ${ }^{[11}$ Furthermore, $\Psi_{f, g}$ carries a canonical partial order, induced by composition of morphisms:

$$
h \prec k \Longleftrightarrow \exists j . j \circ h=k
$$

In this setting, a $\prec$-least member of $\Psi_{f, g}$ determines a (weakly) universal $\Psi$-structure ${ }^{42}$ While we would still prefer a constructive proof, we might hope that some property of $\Psi_{f, g}$, as a (partially ordered) set, will allow us to prove the existence of a $\prec$-least member.

[^72]
## Universal $\Psi$-Structures in $\mathcal{N}$

The class of $\Psi$-structures is a set in $\mathcal{N}$ (as is the equivalence class of universal $\Psi$ structures) ${ }^{43}$ Nevertheless, as with the standard case (i.e. $\mathrm{ZF}(\mathrm{C})$ ), our goal is to produce a witness - a functorial universal $\Psi$-structure - that is a member of the subset describing the semantic property of a universal $\Psi$-structure:

$$
\bar{\Psi}=\{h \mid \forall k \in \Psi . \exists!j . j \circ h=k\} \subset\{x \mid \Psi(x)\}
$$

Typically, we start with the formula $\Phi$, which induced a (functorial) universal structure in Set. If $\Phi$ is stratified and $\{x \mid \Phi(x)\} \in \bar{\Psi}$, the universal structure in Set coincides with that in $\mathcal{N}$.

Remark (Relative Adjoints Cohere with Semantic Universal Structures). If $\Phi$ is unstratified $[4]$ we "type-adjust" the original structure $\Psi$ to a formula $\Psi^{\prime}$ (i.e. a formula that is homogeneous with $\Phi$ ), and study the diagram described by $\Psi^{\prime}$. If $\Phi$ is a member of the set of universal $\Psi^{\prime}$-structures in $\mathcal{N}$, we obtain a relative adjunction. But the set of diagrams of shape $\Psi^{\prime}$ forms a subset of the diagrams of $\Psi$, as the type-adjusted variables will range over subsets $T^{n} V \subset V{ }^{45}$ Thus, $\Phi$ corresponds to a syntactic selection among the set of (semantic) universal $\Psi$-structures, restricted to the subset of diagrams that are also of shape $\Psi^{\prime}$.

The fact that $\Phi_{f, g}$ is not a universal $\Psi_{f, g}$-structure, but is a universal $\Psi_{T f, T g}$-structure, implies that, while $T$ appears to arise as a "syntactic trick," it respects some aspect of the semantic (i.e. category theoretic) structure described by $\Psi$. Therefore, we can ask questions about how properties of $T$ as a functor determine the relationship between syntactic and semantic universal structures. Two natural questions are:

1. Does $T$ preserve/create universal structures?

[^73]2. Is the strength required for the collection of diagrams of shape $\Psi^{\prime}$ to be equivalent to those of shape $\Psi$ (i.e. $\mathcal{N}$ possesses syntactic universal $\Psi$-structures), stronger than what is required for the existence of (semantic) universal $\Psi$-structures?

We can address both questions in the context of coequalizers.

## Coequalizers in NF

The set of coequalizing morphisms of any parallel pair $f, g: Y \rightarrow Z$ is defined:

$$
\Psi_{f, g}=\{h \mid h \circ f=h \circ g\}
$$

A universal coequalizing morphism is a $\prec$-initial member of $\Psi_{f, g}$, where $\underbrace{\sqrt[46]{ }}$

$$
h \prec k \Longleftrightarrow \exists j . j \circ h=k
$$

Each member $h$ of $\Psi_{f, g}$ induces a partition of $Z$ by the equivalence relation:

$$
\hat{h}=\left\{\left\langle z, z^{\prime}\right\rangle \mid h(z)=h\left(z^{\prime}\right)\right\}
$$

We can extend this to each member of $\Psi_{f, g}$ :

$$
\widehat{\Psi}_{f, g}=\{\hat{h} \mid h \circ g=h \circ g\}
$$

We obtain a lower bound on $\widehat{\Psi}_{f, g}$ among the set of equivalence relations on $Y$ :

$$
\downarrow \Psi_{f, g}=\bigcap \widehat{\Psi}_{f, g}
$$

Distinct morphisms inducing equivalent partitions of $Y$ are identified in $\widehat{\Psi}_{f, g}$. Nevertheless, the identification of morphisms with their induced equivalence relations is (weakly) monotonic.

Lemma 3.39. For $h, k \in \Psi_{f, g}, h \prec k \Longrightarrow \hat{h} \subseteq \hat{k}$.
Corollary 3.40. $h$ is a universal coequalizing morphism of $f$ and $g$ if and only if $\hat{h}$ is $a \prec$-initial member of $\widehat{\Psi}_{f, g}$. Furthermore, $\hat{h}=\downarrow \Psi_{f, g}$.

[^74]Alternatively:

Proposition 3.41. $\mathcal{N}$ has (semantic) coequalizers if and only if, for any parallel pair $f$ and $g, \downarrow \Psi_{f, g} \in \Psi_{f, g}$.

The way in which the syntactic coequalizer (in $S e t$ ), $\Phi_{f, g}$, relates to $\Psi_{f, g}$ is now clear. $\sim_{f, g}$ induces the finest partition of $Z$, such that:

$$
z \in\left[z^{\prime}\right]_{\sim} \Longleftrightarrow \exists y \cdot f(y)=z \wedge g(y)=z^{\prime}
$$

In $S$ et (or, for a stratified theory, NFU + Choice), the action $z \rightarrow[z]_{\sim_{f, g}}$ induces the coequalizer $\Phi_{f, g}$.

In NF, $z \mapsto[z]_{\sim_{f, g}}$ is inhomogeneous. Of course, $\sim_{f, g}$ is still a definable (stratified) equivalence relation, but it does not necessarily induce a coequalizing morphism in $\Psi_{f, g}$. If $\sim_{f, g} \in \widehat{\Psi}_{f, g}$, it must be a $\subseteq$-least member.

$$
\sim_{f, g} \subseteq \downarrow \Psi_{f, g}=\bigcap \widehat{\Psi}_{f, g}
$$

It is also true that $\downarrow \Psi_{f, g} \in \widehat{\Psi}_{f, g}$ implies the existence of a coequalizer. By definition, a homogeneous selection function would need to exist ${ }^{[77}$ It is possible, therefore, that some parallel pair $f, g$ exists in $\mathcal{N}$, such that:

$$
\downarrow \Psi_{f, g} \in \widehat{\Psi}_{f, g} \wedge \sim_{f, g} \subsetneq \downarrow \Psi_{f, g}
$$

In other words, the (semantic) coequalizer of $f$ and $g$ may exist in $\mathcal{N}$ and induce a strictly coarser partition than $\sim_{f, g}$.

Remark (Considering KF $+\mathbf{C E}$ ). CE implies the existence of a trivial selection function, for the partition induced by $Y / \sim_{\sim_{f, g}}$. In Lemma 3.25 we proved:

$$
\mathcal{K} \models \mathrm{CE} \Longleftrightarrow \mathcal{K} \models \text { syntactic coequalizers }
$$

One might now ask if there is, in fact, a weaker extension of KF, which is still finitely cocomplete?

[^75]In (the stronger theory) NF, $\Psi_{f, g}$ and $\widehat{\Psi}_{f, g}$ are sets, so we can make this question even more precise. Is there an extension of NF, strictly weaker than NF $+\mathbf{C E}$, such that we can prove:

$$
\forall f, g \cdot \operatorname{par}(f, g) \Longrightarrow \downarrow \Psi_{f, g} \in \widehat{\Psi}_{f, g}
$$

By Proposition 3.41, this would imply the existence of (semantic) coequalizers.

## Interaction of CE and T

We have shown that $\mathbf{C E}$ is precisely equivalent to the completion of the syntactic coequalizer diagram. Thus, an extension of NF that both proves the existence of (semantic) coequalizers and is weaker than $\mathrm{NF}+\mathbf{C E}$, would have to prove the existence of a parallel pair $f, g: Y \rightarrow Z$, with a (semantic) coequalizer $c$ such that:

$$
\sim_{f, g \subsetneq} \subsetneq \hat{c}=\downarrow \Psi_{f, g}
$$

Now consider the parallel pair $T f, T g$. As $T$ is full and faithful, $T c$ is universal among coequalizing morphisms in the image of $T$ :

$$
\widehat{T c}=\bigcap\{\widehat{T h} \mid T h \circ T f=T h \circ T g\}
$$

But $T f$ and $T g$ has a syntactic coequalizer in any extension of NF, corresponding to the relative adjoint $\operatorname{coeq}_{T=} \nmid \Delta$ :

$$
T Z \rightarrow \Phi_{f, g}^{\prime}=\left\{[z]_{\sim_{f, g}} \mid z \in Z\right\} ;\{z\} \mapsto \Phi_{f, g}(z)
$$

Therefore, it is not obvious that $T$ preserves (semantic) coequalizers, beyond those morphisms which are in the image of $T$. In the case that $T$ does preserve the universal $\Psi_{f, g}$ structure, we can show that $\hat{c}$ was syntactic to begin with.

Lemma 3.42. $\widehat{T c}=\sim_{T f, T g} \Longrightarrow \hat{c}=\sim_{f, g}$.

Proof. The equivalence relation on $T Z$, induced by $\{z\} \mapsto \Phi_{f, g}(z)$, is equivalent to $\sim_{T f, T g}$. Furthermore, the canonical isomorphism $\left\langle\{y\},\left\{y^{\prime}\right\}\right\rangle \mapsto\left\{\left\langle y, y^{\prime}\right\rangle\right\}$ implies $\sim_{T f, T g} \cong$ $T \sim_{f, g}{ }^{48}$ By the same argument, $\widehat{T c} \cong T \hat{c}$. Thus $\widehat{T c}=\sim_{T f, T g}$ implies $\hat{c}=\sim_{f, g}$.

[^76]Corollary 3.43. If the (semantic) coequalizer $c$ is preserved under $T, \downarrow \Psi_{f, g}=\sim_{f, g}$.

We can now improve upon Theorem 3.25.
Theorem 3.44. CE holds in $\mathcal{N}$ if and only if $\mathcal{N}$ has (semantic) coequalizers and $T$ preserves them. Equivalently, as $T$ is conservative:

$$
\boldsymbol{C} \boldsymbol{E} \Longleftrightarrow T \text { creates coequalizers }
$$

Remark (A Corollary for Cardinals in $N F$ ). The following result is a consequence of $T$ creating coequalizers, relevant to the set theory of NF.

Corollary 3.45. Define a relation:

$$
B \subseteq W O \times W O=\left\{\left.\langle x, y\rangle| | x\right|_{\sim} \cong|y|_{\sim}\right\}
$$

where $W O$ is the set of well-orderings. If $T$ creates coequalizers, $\pi_{1}, \pi_{2}: B \rightarrow W O$ has a (semantic) coequalizer $c^{\prime}: W O \rightarrow \aleph(T|V|)$, where $\aleph(-)$ denotes Hartog's aleph.

Proof. NO is clearly a coequalizer of $T \pi_{1}, T \pi_{2}$. If $T$ creates coequalizers, then $N O$ must be $T$ of some cardinal. We know $N O=\aleph\left(T^{2}|V|\right)$, so we can describe this cardinal as $T^{-1}(N O)=\aleph(T|V|)$.

### 3.4 Modified-Cartesian Closure

A locally small category is one whose hom-sets (i.e. function spaces) are sets in the sense that they are objects of Set. While function spaces in Set are the "literal" example, the universal property they satisfy can be generalized to arbitrary categories.

Definition 3.46. Given a pair of objects $A, B$ in a category $\mathcal{C}$, an exponential object $B^{A}$ satisfies the following condition, for any object $C$ :

$$
\mathcal{C}(A \times C, B) \cong \mathcal{C}\left(C, B^{A}\right)
$$

If $\mathcal{C}$ satisfies this property for any pair of objects, it is said to be a cartesian closed category.

Definition 3.47. A category $\mathcal{C}$ is cartesian closed if, for each object $A$, the product functor $A \times(-): \mathcal{C} \rightarrow \mathcal{C}$ has a right adjoint $(-)^{A}: \mathcal{C} \rightarrow \mathcal{C}$. In other words, if there is a natural bijection between hom-sets:

$$
\mathcal{C}(A \times B, C) \cong \mathcal{C}\left(B, C^{A}\right)
$$

In Set, the unit and co-unit are defined:

$$
\begin{align*}
& \eta_{A, B}: B \rightarrow(A \times B)^{A}: b \mapsto(a \mapsto\langle a, b\rangle)  \tag{unit}\\
& \varepsilon_{A, C}:\left(A \times C^{A}\right) \rightarrow C:\langle a, f\rangle \mapsto f(a) \tag{co-unit}
\end{align*}
$$

While the condition is obvious for the classical category of $\mathrm{ZF}(\mathrm{C})$ sets, we would hope that any base set theory is cartesian closed. Intuitively, the (more general) universal property of exponential objects should arise from its canonical example: function spaces in the base theory. There are more practical reasons, as well. In particular, we would like to formalize a number of properties for locally small categories (in fibred category theory) over our base category of sets.

Colin McLarty has shown that $\mathcal{N}$ is not a cartesian closed category [40]. From Definition 3.47, it is clear that the component morphisms of the unit and co-unit are inhomogeneous. $\mathcal{N}$ has "literal" function spaces but, generally, they do not possess the universal property of exponential objects ${ }^{49}$ From past experience, however, we know this alone does not disprove the possibility of some cartesian closed model of NF. But McLarty's proof is more general. If $\mathcal{N}$ were cartesian closed, it would be a topos. In this general setting, McLarty defines a version of Cantor's Theorem in the internal language. It is then straightforward to show that a topos with a universal object, one into which each object has a monomorphism, must be the trivial topos.

In KF, there is no universal object (the addition of one would yield NF). So McLarty's proof does not go through. Rather than being inconsistent, we prove that a cartesian closed model of KF is actually a model of full Mac Lane set theory. For this we use a result, due to Forster and Kaye:

[^77]Lemma 3.48. [13] $K F+$ 'every set is strongly cantorian' $=$ Mac.

Theorem 3.49. $K F+$ 'the category of sets is cartesian closed' $=M a c$

Proof. By Lemma 3.48, it is sufficient to prove: if $\mathcal{K}$ is cartesian closed, then every set $A$ is strongly cantorian. The cartesian closure adjunction implies the graph of "curry":

$$
\begin{aligned}
& \text { curry }:((A \times B) \Rightarrow C) \rightarrow(B \Rightarrow(A \Rightarrow C)) \\
& (f:\langle a, b\rangle \mapsto f(a, b)) \mapsto(b \mapsto(a \mapsto f(a, b)))
\end{aligned}
$$

is (locally) a set, for any three objects of $\mathcal{K}$. Thus, local to a given set $A$, the graph of the following map:

$$
f_{1}:(\{\emptyset\} \times\{\emptyset\}) \Rightarrow x \mapsto\{\emptyset\} \Rightarrow(\{\emptyset\} \Rightarrow x)
$$

is a set. We also have a stratified set defining the (local) graph of $f_{2}:\{x\} \mapsto((\{\emptyset\} \times$ $\{\emptyset\}) \Rightarrow x)$, as $\{\emptyset\} \times\{\emptyset\} \cong\{\emptyset\}$. Similarly, the type of $\{\emptyset\} \Rightarrow(\{\emptyset\} \Rightarrow x)$ is equivalent to $\{\{x\}\}$, two types above $x$. We define $f_{3}$ as the function witnessing this fact.

The composition $f_{3} \circ f_{1} \circ f_{2}$ maps $\{x\}$ to $\{\{x\}\}$, for each $x$ in $A$. Thus, $f_{3} \circ f_{1} \circ f_{2}$ is just $\iota \upharpoonright \iota$ " . But this implies that $\iota \upharpoonright A$ is a set, so $A$ is strongly cantorian.

## Stratified Analogue to Cartesian Closure in NF

As with Cantor's Theorem, the best we could hope for in NF is the existence of some stratified analogue to cartesian closure. The stratified versions of the unit and co-unit present themselves $\sqrt[50]{50}$

$$
\begin{aligned}
& \eta_{A, B}^{\prime}: T B \rightarrow(A \times B)^{A}:\{b\} \mapsto(a \mapsto\langle a, b\rangle) \\
& \varepsilon_{A, C}^{\prime}: T A \times C^{A} \rightarrow T C:\langle\{a\}, f\rangle \mapsto\{f(a)\}
\end{aligned}
$$

[^78]Propositions 3.50 and 3.51 prove that these are, respectively, the relative unit and co-unit of the relative adjunctions:

$$
\begin{array}{ll}
(A \times-)_{T} \dashv(A \Rightarrow-) & \left(\eta_{A, B}^{\prime}: T B \rightarrow(A \times B)^{A}\right) \\
(T A \times-) \dashv_{T}(A \Rightarrow-) & \left(\varepsilon_{A, C}^{\prime}: T A \times C^{A} \rightarrow T C\right)
\end{array}
$$

Furthermore, pasting of the relative unit and co-unit yields a symmetric lift, allowing us to recover a form of adjoint symmetry, as expressed by the generalized triangle identity.

Proposition 3.50. For any pair of objects, $A, B$, in $\mathcal{N}$, there is a $(T A \times-)$-co-universal arrow, $\varepsilon_{A, C}^{\prime}$ defined above. Equivalently, there is a relative adjunction, $(T A \times-) \dashv_{T}(A \Rightarrow$ $-)$.

$$
\theta_{B, C}: \mathcal{N}(T A \times B, T C) \cong \mathcal{N}(B,(A \Rightarrow C))
$$

Proof. We prove the existence of a natural isomorphism $\theta: \mathcal{N}(T A \times B, T C) \rightarrow \mathcal{N}\left(B, C^{A}\right)$, defined by the action:

$$
(f: T A \times B \rightarrow T C) \mapsto\left(\bar{f}: B \rightarrow C^{A}: b \mapsto(a \mapsto \bigcup f(\{a\}, b))\right)
$$

Thus, not only is $\bar{f}$ a set, but so is the graph of $\theta$.

To prove naturality, given some $h: B \rightarrow D$, we want to prove that the following square commutes:

$$
\begin{gathered}
\mathcal{N}(T A \times D, T C) \xrightarrow{\theta_{D}} \mathcal{N}\left(D, C^{A}\right) \\
-\circ\left(d_{T A} \times h\right) \downarrow \\
\mathcal{N}(T A \times B, T C) \xrightarrow[\theta_{B}]{\longrightarrow} \mathcal{N}\left(B, C^{A}\right)
\end{gathered}
$$

Consider some $g: T A \times D \rightarrow T C$. The composite $g \circ\left(i d_{T A} \times h\right) \in \mathcal{N}(T A \times B, T C)$ is defined by the action $\langle\{a\}, b\rangle \mapsto g(\{a\}, h(b)) . \quad \theta_{B}\left(g \circ\left(i d_{T A} \times h\right)\right)$ is defined by the action $b \mapsto(a \mapsto \bigcup g(\{a\}, h(b)))$. On the other hand, $\theta_{D}(g)$ defines a map $D \rightarrow C^{A}$, $d \mapsto(a \mapsto \bigcup g(\{a\}, d))$. Precomposing this map with $h$ yields the function defined above, by the action $b \mapsto(a \mapsto \bigcup g(\{a\}, h(b)))$. Thus, $\theta$ is natural.

We next need to prove the existence of an inverse transformation. We define $\theta^{-1}$ by the action:

$$
\left(g: B \rightarrow C^{A}\right) \mapsto(\bar{g}: T A \times B \rightarrow T C:\langle\{a\}, b\rangle \mapsto\{g(b)(a)\})
$$

Naturality is proven by the commutativity of the diagram below, for a morphism $h$ : $B \rightarrow D:$


Given some $g: D \rightarrow C^{A}, \theta_{B}^{-1}(g \circ h)$ is defined as $\langle\{a\}, b\rangle \mapsto\{g(h(b))(a)\} . \theta_{D}^{-1}(g)$ is defined by the action $\langle\{a\}, d\rangle \mapsto\{g(d)(a)\}$, and precomposition with $i d_{T A} \times h$ yields the action defining $\theta_{B}^{-1}(g \circ h)$, above.

Naturality in $C$ follows similarly.

All that remains is to prove that $\theta$ and $\theta^{-1}$ form a two-sided inverse. The calculation is straightforward, as functions are extensional in NF.

The proof of $(A \times-)_{T} \dashv(A \Rightarrow-)$ follows along the lines of Proposition 3.50 .
Proposition 3.51. There exists a relative adjunction $(A \times-){ }_{T} \dashv(A \Rightarrow-)$.

Proof. Given some $f: A \times B \rightarrow C$, there exists a unique morphism $\bar{f}: T B \rightarrow C^{A}$ : $\{b\} \mapsto(a \mapsto f(a, b))$. On the other hand, given some $g: T B \rightarrow C^{A}$, there exists a unique $\bar{g}: A \times B \rightarrow C:\langle a, b\rangle \mapsto g(\{b\})(a)$.

Furthermore, these operations are (external) mutual inverses. The components of the relative unit fit into the expected diagram:


Definition 3.52 (Modified-Cartesian Closure). For a given category $\mathcal{C}$ and endofunctor $T: \mathcal{C} \rightarrow \mathcal{C}$, we refer to the pair of relative adjunctions above as modified-cartesian closure.

In the case of presheaves, the symmetric lift formed by the Yoneda Extension "approximates" a left adjoint. Similarly, the relative unit and co-unit defining modified-cartesian closure approximate left adjoints to exponential functors in $\mathcal{N}$.

Theorem 3.53. In $\mathcal{N}$, the relative adjunctions defining modified-cartesian closure form a symmetric lift.


Proof. Consider an arbitrary object $A$.

$$
\begin{gathered}
\varepsilon_{A \times B}^{\prime} \cdot \eta_{A}^{\prime} \times T B: T A \times T B \rightarrow(A \times B)^{B} \times T B \rightarrow T(A \times B) \\
\langle\{a\},\{b\}\rangle \mapsto\langle(b \mapsto\langle a, b\rangle),\{b\}\rangle \mapsto\{\langle a, b\rangle\}
\end{gathered}
$$

Remark (Uniqueness vs. Internal Bijection). Propositions 3.50 and 3.51 prove the existence of a pair of universal natural transformations corresponding to respective relative adjunctions. The first relative adjunction, $(T A \times-) \dashv_{T}(A \Rightarrow-)$, corresponds to a natural internal bijection:

$$
\mathcal{N}(T A \times B, T C) \cong \mathcal{N}(B, A \Rightarrow C)
$$

Thus, we obtain a relative adjunction, both in the sense that we obtain a universal natural transformation $\varepsilon^{\prime}$ and a natural bijection between hom-sets. In a classical (unstratified) theory, each implies the other. In NF, this is not necessarily the case.
$(A \times-)_{T} \dashv(A \Rightarrow-)$ is a relative adjunction, in the sense that the relative unit:

$$
\eta_{A, B}^{\prime}: T B \rightarrow(A \times B)^{A}:\{b\} \mapsto(a \mapsto\langle a, b\rangle)
$$

is a universal transformation. However, the internal bijection to which $\eta^{\prime}$ corresponds is:

$$
\mathcal{N}(T A \times T B, T C) \cong \mathcal{N}\left(T B, C^{A}\right)
$$

as opposed to what we would expect, externally, from the existence of a $T$-relative unit:

$$
\mathcal{N}(A \times B, C) \cong \mathcal{N}\left(T B, C^{A}\right)
$$

The relative adjunction still exists. We could prove it by the universal property, or by the fact that $T: \mathcal{N} \rightarrow \mathcal{N}$ is an embedding of $\mathcal{N}$ into itself. But we are, in some sense, privileging uniqueness (i.e. the universal property) over (set-theoretic) bijectivity.

Despite their apparent differences, however, the existence of each relative adjunction implies the other ${ }^{51}$

Proposition 3.54. If $T$ preserves exponentials, creates limits and is full and faithful:

$$
(A \times-)_{T} \dashv(A \Rightarrow-) \Longleftrightarrow(T A \times-) \dashv_{T}(A \Rightarrow-)
$$

Proof. We prove this by defining successive chains of natural isomorphisms between hom-sets. The first isomorphism in each chain, however, is "external," in the sense that we do not require it to be internally definable, even if the category has internal function spaces.

$$
\begin{array}{lr}
(\Rightarrow) \\
\operatorname{hom}(B, A \Rightarrow C) & \\
\cong \operatorname{hom}(T B, T(A \Rightarrow C)) & (T \text { is full and faithful }) \\
\cong \operatorname{hom}(T B, T A \Rightarrow T C) & (T \text { preserves exponentials }) \\
\cong \operatorname{hom}(T A \times B, T C) & \left(A \times-{ }_{T} \dashv A \Rightarrow-\right) \\
(\Leftarrow) & (T \text { is full and faithful, and creates limits }) \\
\operatorname{hom}(A \times B, C) & \left(T A \times-\dashv_{T} A \Rightarrow-\right) \\
\cong \operatorname{hom}(T A \times T B, T C) & \\
\cong \operatorname{hom}(T B, A \Rightarrow C) &
\end{array}
$$

[^79]
## Connecting Modified-Cartesian Closure and Powerobjects

Remark (Pseudo-Powerobjects). There are two standard definitions of an elementary topos: a category with finite limits and powerobjects; and a cartesian closed category with finite limits and a subobject classifier. There is another way of stating the existence of powerobjects in topos: ' $\operatorname{Sub}(A \times-)$ is representable, for each object $A$.' In an SPE, the existence of a subobject classifier and modified-cartesian closure implies that $\operatorname{Sub}(T A \times-)$ is representable ${ }^{52]}$ Thus, the stratified membership relation $\in_{A} \subset T A \times A$ is internally definable.

Definition 3.55. We say that a category $\mathcal{C}$ has modified-powerobjects, relative to an endofunctor $T$, if $\operatorname{Sub}(T A \times-)$ is representable for all $A$.

Lemma 3.56. $(T A \times-) \dashv_{T}(A \Rightarrow-)$ implies $\mathcal{N}(T A \times B, \Omega) \cong \mathcal{N}(B, P A)$

Proof. Concrete finite sets are strongly cantorian, so $T \Omega \cong \Omega$.

In [14], we prove the result for an arbitrary SPE with a subobject classifier fixed by $T$.
Lemma 3.57. 14] $(T A \times-) \dashv_{T}(A \Rightarrow-)$ and the existence of a subobject classifer $\Omega$, fixed by $T$, implies the representability of $\operatorname{Sub}(T A \times-)$.

### 3.5 Local Modified-Cartesian Closure

The Fundamental Theorem of Topos Theory states: Given a topos $\mathcal{E}$, each slice category $\mathcal{E} / C$ is a topos, itself. In other words, a category that is a topos globally is a topos locally. Local cartesian closure corresponds to the bottom half of the following adjoint triple, given any morphism $f: C \rightarrow D$ in $\mathcal{E}$ :


[^80]In addition to its role as the "local product," the pullback functor $f^{*}: \mathcal{E} / D \rightarrow \mathcal{E} / C$ corresponds to an internal change of base between indexed families displayed by maps in $\mathcal{E} / D$ and $\mathcal{E} / C: f^{*}:\left(A_{d}\right)_{d \in D} \mapsto\left(A_{f(c)}\right)_{c \in C}$. The adjoint triple describes constructive universal and existential quantification (i.e. dependent products and sums) as adjoint to substitution (i.e. change of base):

$$
\begin{array}{lr}
f^{*}(\alpha: A \rightarrow D) \sim\left(A_{f(c)}\right)_{c \in C} & \text { (change of base) } \\
\Sigma_{f}(\beta: B \rightarrow C) \sim\left(\coprod_{f(c)=d}\left(B_{c}\right)\right)_{d \in D} & \text { (dependent sum) } \\
\Pi_{f}(\beta: B \rightarrow C) \sim\left(\prod_{f(c)=d}\left(B_{c}\right)\right)_{d \in D} & \text { (dependent product) }
\end{array}
$$

In Set, the classical category of sets, the dependent sum and product correspond to:

$$
\begin{aligned}
& \Sigma_{f}(\beta): b \mapsto f \circ \beta(b) \\
& \Pi_{f}(\beta)=\pi_{2}:\left\{\langle g, d\rangle \mid g: f^{-1}(d) \rightarrow A \wedge \beta \circ g=i d_{f^{-1}(d)}\right\} \rightarrow D
\end{aligned}
$$

Therefore, in Set, any indexed family can be displayed internally, along with the the left and right adjoint to re-indexing, $\Sigma$ and $\Pi$. For $\mathcal{N}$, the situation is more complicated ${ }^{53}$ The definition of $\Pi$ is unstratified. As such, the appropriate "fundamental theorem" for $\mathcal{N}$ (and SPE's, in general) describes a pair of relative right adjoints to the pullback functor, which paste to form a symmetric lift.

It does not require proof that, given a morphism $f: C \rightarrow D$ in $\mathcal{N}$, the left adjoint $\Sigma_{f}$ is stratified. Thus, NF can display indexed sums over families indexed by sets of singletons. On the other hand, if a right adjoint to the pullback functor $f^{*}$ were to exist in general (it need not be implemented as $\Pi$ - i.e. that $\Pi$ is unstratified does not, alone, disprove the existence of a right adjoint) we would obtain a contradiction. For an arbitrary object $A$, let $\alpha$ denote the unique morphism to the terminal object. The respective composites $\Sigma_{\alpha} \alpha^{*}$ and $\Pi_{\alpha} \alpha^{*}$ would then produce the standard productexponential adjunction, contradicting McLarty's result. Nevertheless, NF proves the existence of a "best approximation" of dependent products, referred to as modifieddependent products.

[^81]Definition 3.58 (Modified-dependent Products). For a morphism $f: C \rightarrow D$ in $\mathcal{N}$, $\tilde{\Pi}_{f}: \mathcal{N} / C \rightarrow \mathcal{N} / T D$ is defined by the following action on a map $\beta: B \rightarrow C$ in $\mathcal{N} / C$ :

$$
\pi_{2}:\left\{\langle g,\{d\}\rangle \mid g: f^{-1}(d) \rightarrow B \wedge \beta \circ g=i d_{f^{-1}(d)}\right\} \rightarrow T D
$$

There is a relative adjunction $f^{*} T_{D} \dashv \tilde{\Pi}_{f}$, where:

$$
\begin{gathered}
T_{D}: \gamma: G \rightarrow D \mapsto T \gamma: T G \rightarrow T D \\
\mathcal{N} / D \xrightarrow{f^{*}} \mathcal{N} / C \\
T_{D} \downarrow \tilde{\Pi}_{f} \\
\mathcal{N} / T D
\end{gathered}
$$

Furthermore, for $\alpha: A \rightarrow 1$, the composite $\tilde{\Pi}_{\alpha} \alpha^{*}$ yields the exponential functor $A \Rightarrow-$.
Proposition 3.59. Given any morphism $f: C \rightarrow D$ in $\mathcal{N}$, the pullback functor $f^{*}(-)$ is a $T_{D}$-relative left adjoint to $\tilde{\Pi}_{f}(-)$. For any $(\gamma, A)$ in $\mathcal{N} / D$, there is a $\tilde{\Pi}_{f}$-universal arrow $\eta_{\gamma}: T \gamma \rightarrow \tilde{\Pi}_{f}\left(f^{*}(\gamma)\right)$.

Proof. Consider the following diagram:


We prove the above theorem in two parts. First, we prove the existence of an external bijection:

$$
\mathcal{N} / C\left(f^{*}(\gamma), \beta\right) \cong \mathcal{N} / T D\left(T \gamma, \tilde{\Pi}_{f}(\beta)\right)
$$

Second, we prove the existence of a natural transformation $\eta$, between $T_{D}$ and $\tilde{\Pi}_{f} f^{*}$, satisfying the universal property of a $T_{D}$-relative left adjoint.
(1) Any $k:(T A, T \gamma) \rightarrow\left(\tilde{\Pi}_{f}(\beta), \pi_{2}\right)$ in $\mathcal{N} / T D$ is defined by a homogeneous action:

$$
\{a\} \mapsto\langle g, \gamma(a)\rangle
$$

where each element $\{a\} \in T A$ is sent to a section $g$ of $\beta$ restricted to the fibre of $\gamma(a)$ along $f$. Thus, for a given map $k$, we define $\bar{k}:\left(f^{*}(\gamma), \pi_{1}\right) \rightarrow(B, \beta)$ by the action:

$$
\bar{k}:\langle c, a\rangle \mapsto\left(\pi_{1}(k(\{a\}))(c)\right.
$$

For this map to be well-defined, $c$ must be in the fibre $f^{-1}(\gamma(a))$, but this holds, as $\langle c, a\rangle$ is an element of the pullback of $\gamma$ along $f$. Furthermore, $\bar{k}$ is a map over $C$, as $g$ is a section of $\beta$ :


Any $h:\left(f^{*}(\gamma), \pi_{1}\right) \rightarrow(B, \beta)$ is defined by a homogeneous action:

$$
\langle c, a\rangle \mapsto b ; \text { where } \beta(b)=c \wedge f(c)=\gamma(a)
$$

We define the map $\hat{h}:(T A, T \gamma) \rightarrow\left(\tilde{\Pi}_{f}(\beta), \pi_{2}\right)$ by the action:54

$$
\hat{h}:\{a\} \mapsto\langle h(-, a),\{\gamma(a)\}\rangle
$$

We need to confirm that $h(-, a)$ is a section of $\beta$ over $f^{-1}(\gamma(a))$. $h$ has the pullback as its domain, so it is total over the fibre. As $h$ is a map over $C, \beta(h(c, a))=c$. Thus, $h(-, a)$ is a section of $\beta$, as required. As with $\bar{k}$ over $C$, it is straightforward to show that $\hat{h}$ is a morphism in $\mathcal{N} / T D$. By construction, ( $\hat{( })$ and $\overline{()}$ form an external bijection.
(2) We form the map $\eta_{\gamma}: T \gamma \rightarrow \tilde{\Pi}_{f}\left(f^{*}(\gamma)\right)$, by taking the image of $i d$ under the external $\operatorname{map} \hat{( })$, defined in (1). $\widehat{i d}_{f^{*}(\gamma)}$ is defined:

$$
\begin{aligned}
& \widehat{i d}_{f^{*}(\gamma)}:\{a\} \mapsto\langle i d(-, a),\{\gamma(a)\}\rangle \\
& i d(-, a): c \mapsto\langle c, a\rangle
\end{aligned}
$$

Therefore, $\eta_{\gamma}=\overline{i d}_{f^{*}(\gamma)}$ is a stratified definition of a map in $\mathcal{N} / T D$ from $T \gamma$ to $\tilde{\Pi}_{f}\left(f^{*}(\gamma)\right)$. It remains to show the following diagram commutes, for a unique $\bar{k}$.


Given some map $k: T \gamma \rightarrow \tilde{\Pi}_{f}(\beta)$, the unique map $\widehat{k}$ is simply the map $\bar{k}:\langle c, a\rangle \mapsto$ $\pi_{1}(k(\{a\}))(c)$, defined above. The image of $\bar{k}$ under the functor $\tilde{\Pi}_{f}$ is just the map ${ }^{54} h(-, a): c \mapsto h(c, a)$.
$\langle g,\{d\}\rangle \mapsto\langle\bar{k} \circ g,\{d\}\rangle$. On the one hand, $\tilde{\Pi}_{f}(\hat{k}) \circ \eta_{\gamma}$ is defined:

$$
\begin{aligned}
& \tilde{\Pi}_{f}(\hat{k}) \circ \eta_{\gamma}:\{a\} \mapsto\langle i d(-, a),\{\gamma(a)\}\rangle \mapsto\langle\bar{k} \circ i d(-, a),\{\gamma(a)\}\rangle \\
& \bar{k} \circ i d(-, a): c \mapsto\langle c, a\rangle \mapsto \pi_{1}(k(\{a\}))(c)
\end{aligned}
$$

Thus, we conclude $\tilde{\Pi}_{f}(\bar{k}) \circ \eta_{\gamma}=k$.

Naturality of $\eta$ is a straightforward diagram chase.
Corollary 3.60. For any object $A$ and $\alpha: A \rightarrow 1, \tilde{\Pi}_{\alpha} \alpha^{*}(-)=(A \Rightarrow-)$

Proof. The short version: 1 is strongly cantorian, so the stratified analogue of the dependent product is just what it is in the unstratified case. Explicitly, given any object $C$, the pullback of the unique morphism $C \rightarrow 1$ along $\alpha$ is just the product projection $\pi_{1}: A \times C \rightarrow A . \tilde{\Pi}_{\alpha}\left(\pi_{1}\right)$ is the set:

$$
\left\{\langle g, *\rangle \mid g: A \rightarrow A \times C \wedge \pi_{1} \circ g=i d_{A}\right\}
$$

As 1 is strongly cantorian, the set is in bijection with:

$$
\left\{g \mid g: A \rightarrow A \times C \wedge \pi_{1} \circ g=i d_{A}\right\}
$$

The collection of sections of the $\pi_{1}: A \times C \rightarrow A$ is then the collection of functional relations from $A$ to $C$.

Thus, $(A \times-)_{T} \dashv(A \Rightarrow-)$ is equivalent to $\Sigma_{\alpha} \alpha^{*}{ }_{T} \dashv \tilde{\Pi}_{\alpha} \alpha^{*}$.
As $\tilde{\Pi}_{f}$ is a stratified analogue of a dependent product, components $\eta_{\gamma}$ of the relative unit correspond to indexed diagonal maps. Interpreting $\gamma: A \rightarrow D$ as a $T D$-indexed family, $\eta_{\gamma}$ is the $T^{2} D$-indexed family of maps, each of which corresponds to the diagonal map $\delta$ into the indexed product. An explicit example: given some $\alpha: A \rightarrow 1$ and the unique map $!_{2}: 2 \rightarrow 11^{55}, \eta_{\alpha}: A \rightarrow \tilde{\Pi}_{!_{2}}\left(!_{2}^{*}(\alpha)\right)$ is just the pair of $\delta$ maps, $a \mapsto\langle a, a\rangle$.

But asymmetry of the relative adjunction $f^{*} T_{D} \dashv \tilde{\Pi}_{f}$ means we do not obtain a dual universal mapping that gives stratified projections. The stratified projection maps are given by a second relative right adjoint $T f^{*}(-) \dashv_{T_{C}} \tilde{\Pi}_{f}(-)$.

[^82]Proposition 3.61. Given any morphism $f: C \rightarrow D$, the following bijection holds (internally) for any maps $\gamma: A \rightarrow T D$ and $\beta: B \rightarrow C$ :

$$
\mathcal{N} / T C\left(T f^{*}(\gamma), T \beta\right) \cong \mathcal{N} / T D\left(\gamma, \tilde{\Pi}_{f}(\beta)\right)
$$

In other words, there is a relative adjunction $T f^{*}(-) \dashv_{T_{C}} \tilde{\Pi}_{f}(-)$.

Proof. Consistent with the non-localized version of modified-cartesian closure, the bijection defining the relative co-unit is internal. Given some $k: T f^{*}(\gamma) \rightarrow T \beta$, we can form the map $\hat{k}: \gamma \rightarrow \tilde{\Pi}_{f}(\beta)$, defined by the action $a \mapsto\left\langle T^{-1} k(-, a), \gamma(a)\right\rangle$. $k(-, a): T C \rightarrow T B$, so we write $T^{-1} k(-, a)$ to denote the underlying map $C \rightarrow B$, which exists as $T$ is full and faithful. This allows us to "type-lower" and obtain a homogeneous action. Furthermore, $T^{-1} k(-, a)$ is a section of $\beta$, restricted to the fibre over the $d \in D$, such that $\{d\}=\gamma(a)$. In the other direction, given some map $h: \gamma \rightarrow \tilde{\Pi}_{f}(\beta)$, we can form a map $\bar{h}: T f^{*}(\gamma) \rightarrow T \beta$, given by the action $\langle\{c\}, a\rangle \mapsto\left\{\pi_{1}(h(a))(c)\right\}$. Again, this action is homogeneous and one can readily check that both maps form a two-sided inverse, as in the case of the relative unit.

The components of the relative co-unit are given by $\overline{i d}_{\tilde{\Pi}_{f}(\beta)}$. Explicitly, we obtain $\varepsilon_{\beta}$, defined by the action:

$$
\varepsilon_{\beta}:\langle\{c\},\langle g,\{f(c)\}\rangle\rangle \mapsto\{g(c)\}
$$

where $g$ is a section of $\beta$, restricted to $f(c)$. Proving naturality is straightforward.

In the case that $D=1$ (i.e $\left.f=!_{C}: C \rightarrow 1\right), \varepsilon_{\beta}: T f^{*}\left(\tilde{\Pi}_{f}(\beta)\right) \rightarrow T \beta$ is given by stratified evaluation $\langle\{c\}, g\rangle \mapsto\{g(c)\}$, where $g$ is a section of $\beta$. Thus, interpreting $\varepsilon_{\beta}$ as a $T^{2} C$-indexed family of mappings, we obtain the stratified analogue of each projection function $\pi_{\{\{c\}\}}: \Pi_{T C} B_{\{c\}} \rightarrow T B_{\{c\}}$. A "standard" indexed product would require not only that $C$ be cantorian, but that $\beta$ be a map with cantorian fibres. This corresponds to Definition 3.69, where smallness is defined as a fibrewise property of maps, giving us the ability to form (standard) indexed products for "small families of small sets."

The relative unit and co-unit defining local modified-cartesian closure are coherent "approximations" of a left adjoint to $\tilde{\Pi}_{f}$, for each $\operatorname{map} f$ in $\mathcal{N}$.

Theorem 3.62. Given a morphism $f: C \rightarrow D$ in $\mathcal{N}$, the relative unit and co-unit defining local modified-cartesian closure form a symmetric lift.


Proof. For an arbitrary map $\gamma: A \rightarrow D$ :

$$
\varepsilon_{f^{*} \gamma} \circ T f^{*}\left(\iota_{\gamma}\right): \mathcal{N} / D \rightarrow \mathcal{N} / T C
$$

witnesses the isomorphism:

$$
T\left(f^{*}(\gamma)\right) \cong T f^{*}(T \gamma)
$$

Recall the two actions, defined above, which correspond to the relative unit and co-unit:

$$
\begin{aligned}
& \iota_{\gamma}: T \gamma \rightarrow \tilde{\Pi}_{f} f^{*}(\gamma) ;\{a\} \mapsto\langle i d(-, a),\{\gamma(a)\}\rangle \\
& \varepsilon_{f^{*}(\gamma)}: T f^{*} \tilde{\Pi}_{f^{*}(\gamma)} \rightarrow T\left(f^{*}(\gamma)\right) ;\langle\{c\},\langle i d(-, a),\{\gamma(a)\}\rangle\rangle \mapsto\{\langle c, a\rangle\}
\end{aligned}
$$

The image of $\iota_{\gamma}$ under the pullback functor $T f^{*}$ is defined by the action:

$$
\langle\{c\},\{a\}\rangle \mapsto\langle\{c\},\langle i d(-, a),\{\gamma(a)\}\rangle\rangle
$$

Hence the pasting of the relative unit and co-unit is defined by the action:

$$
\langle\{c\},\{a\}\rangle \mapsto\langle\{c\},\langle i d(-, a),\{\gamma(a)\}\rangle\rangle \mapsto\{\langle c, a\rangle\}
$$

Such an action clearly defines an isomorphism in $\mathcal{N} / T C$.


Once again, the isomorphism defining the symmetric lift ultimately depends upon $j^{n}(\iota) \cong$ $j^{n+1}(\iota)$. While straightforward in $\mathcal{N}$, this provides further evidence to support a point about more general SPE's: There appears to be a need for a (canonical) natural isomorphism class of $T$ functors.

### 3.6 Smallness Conditions for NF

Algebraic Set Theory (AST) is centered around the study of categories whose objects behave like proper classes (including a universe object), but have a subcategory of "small" objects, forming an elementary topos [24]. The algebraic aspect of AST involves the construction of "ZF"-Algebras, where models of (I)ZF arise as free initial algebras [24]. Foundation, for example, turns out to be a consequence of Lambek's Lemma [27].

AST should not be expected to transfer to NF in a straightforward manner. Among other things, the set theory of (I)ZF forms a Heyting structure, whereas the sets of NF (under the $\subseteq$-relation) form a Boolean algebra. Nevertheless, there are a number of intriguing connections. It is relatively straightforward to prove that $\mathcal{N}$ satisfies the axioms of a category of classes ${ }^{56}$ and has a strong universe object. Furthermore, the strongly cantorian sets of NF form a topos (Theorem 3.64).

### 3.6.1 Algebraic Set Theory of NF

Definition 3.63 (Class Categories). A class category has four main components:

1. A category $\mathcal{C}$ of classes
2. A subcollection $\mathcal{S} \subset \operatorname{Mor}(\mathcal{C})$ of small maps
3. A powerclass functor $P_{\mathcal{S}}$ that restricts to the powerobject functor on the full subcategory of small objects.
4. A universe object into which each object has a monomorphism $\sqrt{57}$
[^83]An object $C$ is small if its canonical map to the terminal object is a member of $\mathcal{S}$. The full subcategory of small objects is a topos.

In NF, the folklore definition of small is strongly cantorian. Just as the full subcategory of small sets of a class category forms a topos, the full subcategory of $\mathcal{N}$, formed of strongly cantorian sets, forms a topos.

Theorem 3.64. The full subcategory SC of $\mathcal{N}$ of strongly cantorian sets is a topos.

Proof. Recall, a set $A$ is strongly cantorian if $\iota \upharpoonright A$ is a set. Hence, any subset $B$ of $A$ is strongly cantorian, witnessed by $(\iota \upharpoonright A) \upharpoonright B$. Thus, $S C$ is closed under the formation of equalizers. $S C$ is closed under products, $A \times B$, as we can form the isomorphism $\langle\{a\},\{b\}\rangle \mapsto\{\langle a, b\rangle\}$, to obtain the singleton function on $A \times B$. Finally, the strongly cantorian sets are closed under powersets (powerobejects, in $S C$ ). If $A$ is strongly cantorian, the map $S \mapsto(\iota \upharpoonright A) \upharpoonright S: P A \rightarrow P(\iota$ " $A)$ is a set. Composing this with $P(\iota " A) \cong \iota " P A$ gives a singleton function restricted to $P A$. If $S C$ is closed under powersets, then $\operatorname{Sub}(T A \times-)$ is representable for all $A$ in $S C$. But, for all $A$ in $S C$, $T A \cong A .58$

Although NF proves the existence of a natural numbers object $N$ in $\mathcal{N}$, the statement that $N$ is strongly cantorian is strong. Proving that $S C$ is a topos with a natural numbers object, therefore, requires working in an extension of NF.

Corollary 3.65. $N F+A x C o u n t \vdash S C$ is a topos with an NNO.

Definition 3.66. A category $\mathcal{C}$ is a category of classes if the following hold:

1. $\mathcal{C}$ has finite limits.
2. $\mathcal{C}$ has coproducts.
3. $\mathcal{C}$ has kernel quotients, and regular epimorphisms are stable under pullback.

[^84]4. $\mathcal{C}$ has dual images. In other words, for every $f: C \rightarrow D, f^{*}: \operatorname{Sub}(D) \rightarrow \operatorname{Sub}(C)$ has a right adjoint $f_{*}: \operatorname{Sub}(C) \rightarrow \operatorname{Sub}(D)$.

The first and third condition imply that $f^{*}$ has a left adjoint, $f_{!}$, with $f_{!} \dashv f^{*} \dashv f_{*}$ satisfying the Beck-Chevalley conditions. The adjoint triple does not imply cartesian closure, as $f^{*}$ is restricted to subobjects. This reflects the differences between what is desirable for a category of sets and what is desirable for classes.

Proposition 3.67. $\mathcal{N}$ is a category of classes.

Proof. The first three properties were proven in Section 3.2. The final requirement might seem problematic, as pullback functors do not generally have right adjoints in $\mathcal{N}$. For subobject categories, however, $f_{*}$ is the function:

$$
T \in S u b(C) \mapsto\{d \mid \forall c . f(c)=d \Rightarrow c \in T\}
$$

which is clearly homogeneous.

One can now approach the question of small sets in NF by investigating the properties of small objects in a class category. In the categorical setting, smallness is defined fibre-wise.

Definition 3.68. A subcollection $\mathcal{S}$ of $\operatorname{Mor}(\mathcal{C})$, for a category of classes $\mathcal{C}$, is said to be a system of small maps if it satisfies the following conditions:

1. $\mathcal{S}$ is closed under composition and contains all identity morphisms.
2. The pullback of a small map along any map is small.
3. Diagonal $\Delta: C \times C \rightarrow C$ are small.
4. If $f \circ e$ is small and $e$ is a regular epimorphism, then $f$ is small.
5. Copairs of small maps are small.

Closure under arbitrary pullbacks is a requirement of a "fibre-wise" condition. ' $\mathcal{S}$ has diagonals' is equivalent to saying 'the equivalence relation under equality is small.' The fourth condition is less obvious, but essentially says: covering maps respect smallness. The fifth condition simply asserts that smallness is stable under pairing.

Definition 3.69. A map $f: C \rightarrow D$ in $\mathcal{N}$ is said to be strongly cantorian if each fibre $f^{-1}(d)$ is a strongly cantorian set.

The first axiom of a system of small maps may seem the most banal - in particular, the closure of $\mathcal{S}$ under composition. Behind this condition, however, is an important general principle of "smallness." As small morphisms correspond to maps with small fibres, preservation under composition requires: the sum-set of a small family of small sets is small. It does not appear to be a theorem of NF that strongly cantorian sets satisfy this condition.

One solution is the existence of a choice principle, but this is tricky business for NF. It is preferable to consider an extension of NF, much as one does with the Axiom of Counting, which directly adds the sum-set condition to the axiom scheme.

Definition 3.70 (SCU Axiom). The sum-set of a strongly cantorian set of strongly cantorian sets is strongly cantorian.

Proposition 3.71. In $N F+S C U$, the collection of strongly cantorian maps, $\boldsymbol{S C}$, forms a system of small maps.

Proof. In the presence of SCU, SC is closed under composition and contains identity morphism. Given a strongly cantorian map, $f: C \rightarrow D$, and an arbitrary map, $g$ : $B \rightarrow D$, consider $g^{*}(f)=\pi_{1}: B \times_{D} C \rightarrow B$. The fibre over an arbitrary $b \in B$ is the collection of pairs $\langle b, c\rangle$ such that $g(b)=f(c)$, thus it is in bijection with $f^{-1}(g(b))$ and, therefore, strongly cantorian. Copairs of strongly cantorian maps and diagonal functions are clearly strongly canorian. It remains to prove that if the composite $f \circ e$ is strongly cantorian, and $e$ is a regular epimorphism, $f$ is strongly cantorian. Given any object $d$ in the domain of $f \circ e,\left|f^{-1}(d)\right| \preceq\left|(f \circ e)^{-1}(e(d))\right|$. It is a result of [9] that:

$$
\operatorname{stcan}(x) \wedge|y| \preceq \operatorname{stcan}(x) \Longrightarrow \operatorname{stcan}(y)
$$

so each fibre of $f$ is strongly cantorian.

Corollary 3.72. The collection of strongly cantorian maps, $\boldsymbol{S C}$, also satisfies a "descent condition." Given a strongly cantorian map, $f$, and a regular epimorphism, e, fitting into a pullback diagram below, $g$ is strongly cantorian:


Proof. Consider any fibre $g^{-1}(a)$. As $e$ is surjective, there is some $c \in C$ such that $e(c)=a$. As $g^{-1}(a)$ injects into the strongly cantorian fibre $f^{-1}(c)$, it must itself be strongly cantorian. Hence, $g$ is a strongly cantorian map.

While $S C U$ is not apparently provable in $N F$, it is a theorem of $N F U+$ Choice.

Lemma 3.73. $N F U+$ Choice $\vdash S C U$

Proof. Let $X$ be a strongly cantorian set of strongly cantorian sets. AC implies that every strongly cantorian set is the same size as an initial segment of the ordinals (and all the ordinals in that initial segment will be cantorian). Use choice to pick one such bijection for each $x \in X$, and consider the implied action of these on $X$. Thus $\bigcup X$ has an "address" that is an ordered pair of cantorian ordinals, so $\bigcup X$ injects into a set of ordered pairs of cantorian ordinals. Any such set witnesses its cantorian-ness, so must itself be strongly cantorian. So $\bigcup X$ is strongly cantorian, as desired.

Corollary 3.74. In both NFU + Choice and $N F+S C U$, the category of sets is a category of classes, and the collection of strongly cantorian maps is a system of small maps.

If strongly cantorian is truly a good notion of smallness, we should be able to form a powerclass functor $P_{S}$, which takes a set and returns the set of strongly cantorian subsets. Such a functor would restrict to the powerobject functor on $S C$.

Definition 3.75. A relation $R \hookrightarrow A \times B$ is small if its projection onto $B$ is a small map. Thus, a subobject $D$ of $C$ is small, if the projection of $D \hookrightarrow C \times 1$ onto 1 is a small map.

In other words, a relation is small if the set of objects related to any object in the codomain is small. We extend this intuition to the special case of subobject categories. A monomorphism is said to be small as a subobject if its domain is a small. $5^{59}$ In $N F+S C U$, we obtain the easy lemma:

Lemma 3.76. In $N F+S C U$, the calculus of small relations is closed under relational composition.

Proof. Given small relations $R \mapsto A \times B$ and $S \mapsto B \times C$, consider the composite $R \circ S \mapsto A \times C$. For each $c \in C$ :

$$
(R \circ S)^{-1}(c)=\{a \mid \exists b \cdot a R b \wedge b S c\}=\bigcup_{b S c}\{a \mid a R b\}
$$

Thus, $(R \circ S)^{-1}(c)$ is equivalent to a strongly cantorian union of strongly cantorian sets.

Definition 3.77. A powerclass functor $P_{S}$ for a class category $\mathcal{C}$ satisfies the following universal property: For every object $C$ and every small relation $R \hookrightarrow C \times X$, there is a unique arrow $\rho: X \rightarrow P_{S}(C)$ satisfying the following pullback diagram:


Furthermore, the subset relation $\subseteq_{C} \hookrightarrow P_{S}(C) \times P_{S}(C)$ is small.

[^85]At first, with respect to $\mathbf{S C}$, there is reason to be optimistic. The relation $\epsilon_{C} \subset C \times$ $P_{S}(C)$, while not generally stratifiable for the full powerobject functor $P$, is stratifiable for $P_{S}(C)$, where the set of strongly cantorian subsets of $C$ exists. One simply uses the
 by a stratified formula, so the obvious choice of powerclass functor is not definable. In fact, we can show that no functor can exist that satisfies the universal property of a powerclass functor, for the system of small maps, SC.

Proposition 3.78. $N F+S C U$ cannot form a powerclass functor for the system of strongly cantorian maps.

Proof. The image of the universal set $P_{S}(V)$ would be the set of strongly cantorian sets. $P_{S}(V) \cap W O$ would then be the set of strongly cantorian well-orders. As each ordering in the set is a well-ordering of a strongly cantorian set, given an ordering $\leq, \operatorname{RUSC}(\leq)$ and $\leq$ are of the same order type. From this we can conclude, by transfinite induction, that every ordinal number of a strongly cantorian ordinal is the order type of all ordinals beneath it. Then, using the natural ordering on a set of ordinals, the collection of strongly cantorian well-orders will itself be a strongly cantorian well-order. Thus, the collection of strongly cantorian well-orders must be contained in $P_{S}(V) \cap W O$, despite also being longer than any element of $P_{S}(V) \cap W O$ - an instance of the Burali-Forti paradox.

Remark (Strongly Cantorian Indexing). Strongly cantorian is a good notion of external smallness for NF, but not sufficient for internal smallness ${ }^{60}$

Strongly cantorian maps are, nevertheless, an important subclass of $\operatorname{Mor}(\mathcal{N})$, even in the absence of SCU. If we examine the relative adjoints defining modified-dependent products, we see that strongly cantorian products of strongly cantorian sets exist, internally, in NF. Moreover, they arise as strongly cantorian maps over strongly cantorian sets. Thus, strongly cantorian maps allow one to recover a "natural" notion of indexed products from the relative adjunction, $T f^{*}(-) \dashv_{T_{C}} \tilde{\Pi}_{f}(\beta)$.

[^86]Proposition 3.79 (Informal). Small maps over small objects, in $\mathcal{N}$, have natural behavior as indexed families.

### 3.6.2 Consequences of SCU

The extension of $N F$ to $N F+S C U$ was intended to form a set theory, where strongly cantorian was an appropriate category-theoretic notion of smallness. But it also determines a potentially interesting extension of NF from the perspective of set theory. In particular, $N F+S C U$ satisfies desirable smallness conditions, which do not appear to be theorems of the axioms of $N F$.

Knowing that $S C U$ is a theorem of $N F U+$ Choice, we consider possible methods of proving $S C U$ in NF. A common strategy, introduced by Dana Scott, is to consider permutation models [54].

Definition 3.80. If $\langle V, R\rangle$ is a model in the language of set theory, and $\pi$ is a permutation of $V,\left\langle V, R_{\pi}\right\rangle$ is a Rieger-Bernays permutation model, where:

$$
x R_{\pi} y \Longleftrightarrow x R \pi(y)
$$

Given a formula $\varphi$, in the language of set theory, and a permutation $\pi, \varphi^{\pi}$ is the result of replacing each occurrence of ' $x \in y$ ' with $x \in \pi(y)$. As a result:

$$
\langle V, R\rangle \models \varphi \Longleftrightarrow\left\langle V, R_{\pi}\right\rangle \models \varphi^{\pi}
$$

One says that $\varphi$ is invariant if: $\varphi \Longleftrightarrow \varphi^{\pi}$ for all setlike permutations. It is a known result that stratified formulae are invariant. There is, in fact, a stronger result: the stratified sentences of a theory are precisely those that are preserved by each set-like permutation of the universe [9]. Thus, if we have a model of NF, we can permute the universe and obtain a new model in which all stratified sentences in the old model are preserved, but some new formula may be satisfied.

Scott used an argument of this nature to prove that, given any model of NF, there exists a permutation model of NF with Quine atoms. We might hope that such a model can
be used to prove $S C U$, as the formula expressing it is unstratified. Unfortunately, $S C U$ is invariant.

Definition 3.81. Given a setlike permutation $\pi$, we define $\pi_{0}=i d$ and $\pi_{n+1}=\left(j^{n}(\pi)\right) \pi_{n}$.

Lemma 3.82 ([17]). Let $\phi$ be stratifiable with free variables $x_{1}, \ldots, x_{n}$, where $x_{i}$ has been assigned $k_{i}$ in a some valid stratification. Let $\pi$ be a setlike permutation and $V$ a model of NF. Then

$$
(\forall \vec{x}) V \models\left(\phi(\vec{x})^{\pi} \Longleftrightarrow \phi\left(\pi_{k_{1}}\left(x_{1}\right), \ldots, \pi_{k_{n}}\left(x_{n}\right)\right)\right.
$$

Henson's lemma can also be applied to unstratified formula such as $S C U$, as it is with more common axioms (for NF), like the Axiom of Counting.

Lemma 3.83. $S C U$ is invariant.

Proof. In primitive notation, $S C U$ is written:

$$
\forall x .(\operatorname{stcan}(x) \wedge(\forall y)(y \in x \rightarrow \operatorname{stcan}(y))) \rightarrow(\forall z)(z=\bigcup x \rightarrow \operatorname{stcan}(z))
$$

$\operatorname{stcan}(x)^{\sigma}$ is $\operatorname{stcan}(\sigma(x))$ or, equivalently, by Henson's Lemma, stcan $(\sigma " \sigma(x))$. Likewise, $(z=\bigcup x)^{\sigma}$ is $\sigma(z)=\bigcup \sigma^{"} \sigma(x)$. Using these identities, we can write $S C U^{\sigma}$ as:
$(\forall x)(\operatorname{stcan}(\sigma(x)) \wedge(\forall y)(y \in \sigma(x) \rightarrow \operatorname{stcan}(\sigma(y)))) \rightarrow(\forall z)(\sigma(z)=\bigcup \sigma " \sigma(s) \rightarrow \operatorname{stcan}(z))$
The next move is to use the fact that $x$ and $z$ are universally quantified and $\sigma$ is a permutation of the universe. Thus, we can "reletter" ' $x$ ' and ' $z$,' and simplify the formula to:

$$
(\forall x)\left(\operatorname{stcan}(\sigma(x)) \wedge(\forall y)(y \in x \rightarrow \operatorname{stcan}(\sigma(y))) . \rightarrow \operatorname{stcan}\left(\bigcup \sigma^{\prime \prime} x\right)\right)
$$

But, again recalling Henson's Lemma, this is equivalent to:

$$
(\forall x)(\operatorname{stcan}(x) \wedge(\forall y)(y \in x \rightarrow \operatorname{stcan}(y)) . \rightarrow \operatorname{stcan}(\bigcup x))
$$

Thus, we have recovered $S C U$, by a chain of equivalences.
$S C U$ turns out to be equivalent to another property of NF.

Definition 3.84. $\iota_{1}$ is the function that sends each strongly cantorian set $x$ to $\iota \upharpoonright x$.

When $\iota_{1}$ is restricted to the strongly cantorian sets of NF, it is homogeneous. $\iota_{1}$ cannot, however, be defined as a set. Consider $\iota_{1}$ " $(\iota$ " $V)=\{\iota \upharpoonright\{x\} \mid x \in V\}$. If this were a set, we could form the set $\bigcup \iota_{1}$ " $(\iota$ " $V)$, which would be the graph of the singleton function, allowing us to prove Cantor's paradox. It may, however, be the case that $\iota_{1} \upharpoonright x$ is as set when $x$ is strongly cantorian.

Proposition 3.85. SCU is equivalent to the assertion that for all strongly cantorian sets of strongly cantorian sets, $\iota_{1} \upharpoonright x$ is a set.

Proof. First, assume $S C U$, and let $X$ be a strongly cantorian set of strongly cantorian sets. The functional relation $\iota \upharpoonright \bigcup X$ is, therefore, a set. Restricting $\iota \upharpoonright \bigcup X$ to a given $x \in X$ allows us to form a stratified set abstract, as each $x \in X$ is strongly cantorian. But this is equivalent to saying $\iota_{1} \upharpoonright X$ is a set.

Now assume $\iota_{1} \upharpoonright X$ is a set for any strongly cantorian set, $X$, of strongly cantorian sets. As $X$ is a set, the image of $X$ under $\iota_{1}$ is a set $\{\iota \upharpoonright x \mid x \in X\}$. Then $\bigcup\{\iota \upharpoonright x \mid x \in X\}$ is a set. This defines the set $\iota \upharpoonright \bigcup X$, which is equivalent to $S C U$.

SCU implies $\iota_{1} \upharpoonright x$ is a set, for any strongly cantorian set of strongly cantorian sets. NF is not a well-founded theory, but we can consider the concept of hereditarily strongly cantorian sets. In this direction, consider the function $\iota_{2}: x \mapsto \iota_{1} \upharpoonright x$ for each strongly cantorian set of strongly cantorian sets.

Definition 3.86. The notation $\operatorname{stcan}_{2}(x)$ indicates $x$ is a strongly cantorian set of strongly cantorian sets. Generally, $\operatorname{stcan}_{n}(x)$ indicates a set $x$ is hereditarily strongly cantorian going $n$ level down in "rank."

We can then define an axiom scheme:

$$
\begin{align*}
& \iota_{1} \upharpoonright x \text { is a set } \Longleftrightarrow \operatorname{stcan}_{2}(x)  \tag{1}\\
& \iota_{2} \upharpoonright x \text { is a set } \Longleftrightarrow \operatorname{stcan}_{3}(x)  \tag{2}\\
& \ldots  \tag{n}\\
& \iota_{n} \upharpoonright x \text { is a set } \Longleftrightarrow \operatorname{stcan}_{n+1}(x)
\end{align*}
$$

The instances of which turn out to be mutually equivalent.
Proposition 3.87. $\forall n . S C U_{n} \Longleftrightarrow S C U_{n+1}$.

Proof. First, we prove that $S C U_{n+1}$ implies $S C U_{n}$, generally. Suppose $\operatorname{stcan}_{n+1}(x)$, we want to show that $\iota_{n} \upharpoonright x$ is a set. $\operatorname{stcan}_{n+2}(\iota$ " $x)$, so by $S C U_{n+1} \iota_{n+1} \upharpoonright \iota^{\prime \prime} x$ is a set. This takes a value $\{y\} \in \iota$ " $x$ and returns $\iota_{n} \upharpoonright\{y\}$, which is the singleton $\left\{\left\langle y, \iota_{n}(y)\right\rangle\right\}$. Hence, the sumset of $\iota_{n+1}$ " $(\iota x), \bigcup\left\{\left\{\left\langle y, \iota_{n}(y)\right\rangle\right\} \mid y \in x\right\}$ is a set. This is precisely $\iota_{n} \upharpoonright x$, as desired.

In the other direction, assume $S C U_{n}$. It is self-evident that $\operatorname{stcan}_{n+2}(x) \Rightarrow \operatorname{stcan}_{n+1}(\bigcup x)$. $S C U_{n}$ implies that $\iota_{n} \upharpoonright(\bigcup x)$ is a set and can be restricted to any subset of $x$ and, hence, extended to a function on $P(\bigcup(x))$. Strongly cantorian sets are closed under powerset and subset. Therefore, as $P(\bigcup x)$ is a superset of $x, \iota_{n+1} \upharpoonright x$ is a set, proving $S C U_{n+1}$

We are not able to show that $S C U$ is equivalent to the assertion that the transitive closure of a hereditarily strongly cantorian set is strongly cantorian - though it is for any concrete set of finite rank. It seems likely one cannot prove the result in full generality (nor is it obvious it is even true).
$N F+S C U$ allows us to expand the properties of strongly cantorian sets, summarized in [9], to include closure under indexed products and that $\mathbf{S C}$ is closed under directed, strongly cantorian limits.

Theorem 3.88 (SCU). For any strongly cantorian set I, an I-indexed product of strongly cantorian sets, $\Pi_{i \in I} A_{i}$, is strongly cantorian.

Proof. The indexed product is a subset of $P\left(\bigcup_{i \in I} A_{i} \times I\right)$. By $S C U, \bigcup_{i \in I} A_{i}$ is strongly cantorian. We know that a binary product of strongly cantorian sets is strongly cantorian, and that strongly cantorian sets are closed under the powerset operation. Thus, $P\left(\bigcup_{i \in I} A_{i} \times I\right)$ is strongly cantorian. Therefore, $\Pi_{i \in I} A_{i}$ is a subset of a strongly cantorian set.

Theorem 3.89 (SCU). Let $\left\langle I, \leq_{I}\right\rangle$ be a directed, strongly cantorian poset, and let $\left\{A_{i} \mid i \in I\right\}$ be a family of sets with strongly cantorian surjections $\pi_{i, j}: A_{i} \rightarrow A_{j}$, whenever $i>_{I} j$, with all projections commuting. The limit object $A_{I}$ of this directed family is such that each map $\pi_{I, i}: A_{I} \rightarrow A_{i}$ is strongly cantorian (i.e. has strongly cantorian fibres).

Proof. The projective limit $A_{I}$ is defined as:

$$
\left\{f \in \Pi_{i \in I} A_{i} \mid \forall j>_{I} i . \pi_{j, i}(f(j))=f(i)\right\}
$$

For $x \in A_{i}$, the fibre $\pi_{I, i}^{-1 "}\{x\}$ is

$$
\left\{f \in \Pi_{i \in I} A_{i} \mid \forall j>_{I} i \cdot \pi_{j, i}(f(j))=x\right\}
$$

In other words, the fibre over any $x \in A_{i}$ is the set of functions $f$ that pick elements $y_{j}$ in $A_{j}$, for all $j>i$, such that $\pi_{j, i}\left(y_{j}\right)=x$. Thus, it is a subset of the indexed product of the fibres $\pi_{i, j}^{-1}$ " $\{x\}$. As any chain in $I$ is strongly cantorian, and the fibres of $\pi_{j, i}$ are strongly cantorian by assumption, each $\pi_{I, i}$ is strongly cantorian.

Without the Axiom of Counting, one cannot prove that NF contains an infinite strongly cantorian set. But, with this in hand, we can prove that SC has coequalizers and is closed under them. This requires a lemma stating that "locally" strongly cantorian graphs are "globally" strongly cantorian.

Lemma 3.90 (SCU). If $G$ is a connected graph and, for every element $x$, the set $N(x)$ of neighbors of $x$ is strongly cantorian, then the edge and vertex sets of $G$ are both strongly cantorian.

Proof. Given some vertex $v$, consider the countable sequences $\left\langle N_{n}(v) \mid n \in N\right\rangle$, where $N_{n}(v)$ denotes the set of vertices at most $n$ edges away from $v$. Because we do not have unstratified induction, we cannot proceed as one might expect, to prove $\operatorname{stcan}\left(N_{n}(v)\right)$ for all $n$. Using weakly stratified induction over the naturals, however, we can prove that $\iota \cap\left(N_{T n}(v) \times \iota\right.$ " $\left.N_{n}(v)\right)$ exists. The base case of $n=1$ holds, as the set of neighbors of $v$ is strongly cantorian. Now assume that $\iota \cap\left(N_{T n}(v) \times \iota\right.$ " $\left.N_{n}(v)\right)$ exists for some $n \in N$. By the Axiom of Counting, $N_{T n}(v)=N_{n}(v)$, so the restriction of the singleton function in the hypothesis is just $\iota \upharpoonright N_{n}(v)$. For the induction step, we form $N_{n+1}(v)$ as the union of $N_{1}(x)$, for all $x \in N_{n}(v)$. By $S C U$ this is strongly cantorian, so the proof by (weakly stratified) induction is complete. Thus, for all $n \in N, \iota \cap\left(N_{T n}(v) \times \iota\right.$ " $\left.N_{n}(v)\right)$ exists. By Counting, $\iota \upharpoonright N_{n}(v)$ exists for all $n$. But this implies that the vertex set of $G$ is strongly cantorian, as $G$ is connected. Using $S C U$ once more, we obtain that the edge set of $G$ is strongly cantorian, as well.

Theorem 3.91. Given a parallel pair of strongly cantorian maps $f, g: A \rightarrow B$, their (strongly syntactic) coequalizer exists and, furthermore, its universal map is strongly cantorian.

Proof. We take the quotient of the least equivalence relation on $B$, such that $b_{1} \sim b_{2}$ if $\exists a . f(a)=b_{1} \wedge g(a)=b_{2}$. Each element of this quotient can be viewed as a connected graph, with $N(b)$ strongly cantorian, for each $b$ in the equivalence class. Thus, the equivalence class is itself strongly cantorian, and the quotient map is both stratified and strongly cantorian.

## Chapter 4

## Category Theory in NF

The current chapter focuses on what might be phrased "category theory in NF," the study of the theory of categories, taking $\mathcal{N}$ as a universe. The underlying axioms of category theory can be expressed diagrammatically, within a category of limited strength. For example, only assuming finite limits, we are able to interpret internal theories of categories, group(oid)s, presheaves, etc. Indeed, this modest property is sufficient for a reasonable base category, in the richer theory of Fibred Categories [3].

The general study of abstract category theory within arbitrary categories - the study of categories as mathematical universes - is referred to as internal category theory. For Set, internal category theory largely reduces to studying the standard category of small categories. In an arbitrary category, however, one is restricted to those formulas which are expressible in the internal language. In this sense, one is limited by both size (i.e. restricted quantification) and the properties of the category, which determine the strength of the internal language.

The first two sections of this chapter provide an introduction to some of the basic constructions and ideas of internal category theory. In the first section, we give a brief motivation for internal category theory and define the basic constructs. Where there is a meaningful distinction to be made, we contrast the basic internal category theory of Set with that of $\mathcal{N}$. The second section reviews the internal Yoneda Lemma and the
(co)algebraic presentation of internal presheaves, largely in preparation for the following section, where we prove the existence of an NF-Yoneda Lemma. ${ }^{1}$

The (co)algebraic presentation of internal presheaves in $\mathcal{N}$ turns out to be consistent with earlier results on the existence of dependent sums and modified-dependent products. The monadic presentation of internal presheaves holds, but formation of the relevant comonad, in the dual presentation, is unstratified. Using modified-dependent products and the relative version of Eilenberg and Moore's comonadicity theorem, presented in Section 2.4, we are able to prove that the relevant dual structure in $\mathcal{N}$ is a relative comonad.

The existence of apparently non-equivalent internal presentations of presheaves in $\mathcal{N}$ merits investigation. In particular, we note the role of comonadicty in Lawvere and Tierney's fundamental result that categories of internal presheaves in an elementary topos are themselves toposes. We consider an analogue of this result in $\mathcal{N}$, using relative comonadicity, in our effort to understand preservation properties of modified-cartesian closure in categories of (internal) presheaves. This also ties into our consideration of whether the Yoneda embedding is the appropriate $T$ functor for categories of internal presheaves in $\mathcal{N}$.

Externalization and Fam (coproduct completion) are examined in Section 4.4. Here, we encounter higher dimensional relative structures for the first time. We prove that the internalization of Fam in $\mathcal{N}$ is a relative pseudomonad Classically, FamC corresponds to the free coproduct completion of $\mathcal{C}$. In this role, $\operatorname{Fam\mathcal {C}}$ is the canonical example of a KZ-doctrine $\sqrt[3]{3}$ In $\mathcal{N}$, the relative pseudomonad corresponding to the internalization of Fam turns out to be a lax idemptotent relative pseudomonad. This result speaks in two directions. First, it speaks to the appropriate definition of coproduct completeness

[^87]for internal categories in NF: co-completeness with respect to $T$-indexed families. Second, we obtain a specific case of a relative lax idemptotent pseudomonad arising as a coproduct completion for the class of coproducts indexed by objects in the image of the relative functor.

Algebras of KZ-doctrines are uniquely determined by their adjoint relationship with the unit. In the strict case, considered in [26], they arise as left adjoint to unit components of the pseudomonad. In fact, this classification holds generally [36]. The definition of a relative lax idempotent pseudomonad in [8] generalizes [37], which classifies KZ-Doctrines and their algebras as no-iteration pseudomonads, whose algebras are extensions along components of the unit $\left\{^{4}\right.$ In their work introducing relative lax idempotent pseudomonads, Hyland et al. required the Kleisli presentation of relative KZ-algebras [8]. For our purpose, which is to determine whether the internal Fam-algebras are precisely the $T$ coproduct complete internal categories of $\mathcal{N}$, we need to examine the Eilenberg-Moore algebras. This presents a challenge, as it is not clear that the classification theorem for the category of KZ-algebras transfers to the relative case, as it does if we restrict attention to the free (i.e. Kleisli) algebras. Alongside our consideration of the special case of Fam-algebras in $\mathcal{N}$, we seek to make these issues precise in the general case. Referring back to our consideration of "free" relative structures as, in some sense, syntactic, we consider whether the existence of this classification, in the relative case, is a question of semantics (i.e. dependent upon properties of the structure and category in which one is working).

The final section examines the properties of $\tilde{\mathcal{N}}$, the self-internalization of $\mathcal{N}$. We incorporate the fibred notion of local smallness and prove that, despite only having display maps for T-indexed families, the codomain fibration of $\mathcal{N}$ over itself has an object in the fibre over $V$, which gives rise to a full internal subcategory equivalent to $\tilde{\mathcal{N}}$. Such a development allows us to consider more general "universe" structures within $\mathcal{N}$, an interest partially motivated by dependent type theory [67].

[^88]
### 4.1 Basic Internal Category Theory

Section 4.1 proves some basic results regarding internal category theory in $\mathcal{N}$, but is largely a review of elementary internal category theory. Our review follows [22] closely. A more elementary introduction can be found in [5]. For a more advanced and extensive account, bridging internal category theory and indexed/fibred category theory, [23] is as good a reference as any.

## The Basic Definitions

Across this chapter we use the convention of $\mathbb{C}$ for internal categories and $\mathcal{C}$ for standard categories. The one exception is the internal category of NF sets, which we continue to denote $\tilde{\mathcal{N}}$.

Definition 4.1. Given a category $\mathcal{E}$ with finite limits, the 2-category of internal categories, $\operatorname{cat}(\mathcal{E})$, is defined by the following data:

Objects: An internal category $\mathbb{C}$ is a diagram in $\mathcal{E}$

$$
C_{1} \times_{d_{0}} C_{1} \xrightarrow{m} C_{1} \underset{d_{1}}{\stackrel{d_{0}}{\leftarrow}} C_{0}
$$

with equations satisfying the axioms of an elementary category

$$
\begin{align*}
& d_{0} \circ i=d_{1} \circ i=1_{C_{0}}  \tag{unit1}\\
& m \circ\left\langle 1, i \circ d_{1}\right\rangle=m \circ\left\langle i \circ d_{0}, 1\right\rangle  \tag{unit2}\\
& d_{1} \circ m=d_{1} \circ \pi_{2} \wedge d_{0} \circ m=d_{0} \circ \pi_{1}  \tag{comp}\\
& m \circ(1 \times m)=m \circ(m \times 1) \tag{assoc}
\end{align*}
$$

1-cells: An internal functor $F: \mathbb{C} \rightarrow \mathbb{D}$ is a parallel pair of morphisms $F_{0}: C_{0} \rightarrow D_{0}$ and $F_{1}: C_{1} \rightarrow D_{1}$, which commute with the structure maps of $\mathbb{C}$ and $\mathbb{D}$

$$
\begin{array}{lr}
F_{1} \circ i_{C}=i_{D} \circ F_{1} & \left(F\left(1_{c}\right)=1_{F c}\right) \\
F_{1} \circ m_{C}=m_{D} \circ F_{1} \times F_{1} & (F(g f)=F(g) F(f))
\end{array}
$$

2-cells: Given a pair of internal functors $F, G: \mathbb{C} \rightarrow \mathbb{D}$, an internal natural transformation $\tau: F \rightarrow G$ is a morphism $\tau: C_{0} \rightarrow D_{1}$, such that the following equation holds, expressing the naturality condition:

$$
m_{D} \circ\left\langle G_{1}, \tau \circ d_{0}\right\rangle=m_{D} \circ\left\langle\tau \circ d_{1}, F_{1}\right\rangle
$$

We provide two motivating examples:
Example 4.2 (Cat). The category of small categories arises as the category of internal categories in Set, $\operatorname{cat}(S e t)$. Any functor $F$ between small categories is defined by $F_{0}$ and $F_{1}$, its operation on objects and morphisms respectively. In turn, the graphs of $F_{0}$ and $F_{1}$ are morphisms in Set. Not only can Cat be defined as a category of internal categories but, as we will see below, (co)completeness can also be classified internally. Absent sufficient choice functions, however, we have to draw some distinction between category theory as we recognize it (externally) and working purely internal to Set. To avoid this distinction, one can work, as advocated by Makkai, with (internal) anafunctors, where saturated anafunctors correspond to universal properties [34.

Example 4.3 (Topological Groupoids). Top, the category of topological spaces and continuous maps between them, is finitely complete and, therefore, has internal categories whose structure maps are continuous mappings between topological spaces of objects, morphisms, and composable pairs of morphisms. 5 A particularly useful case is the classification of topological groupoids. An internal category $\mathbb{C}$ is an internal groupoid if there exists a twist-isomorphism $\tau: C_{1} \rightarrow C_{1}$, mapping each morphism to its inverse. The morphisms of $\mathbb{C}$ (i.e. the "elements" of $C_{1}$ ) form a canonical groupoid. Indeed, the subcategory $\operatorname{Grpd}($ Top $)$ of internal Top-categories is precisely the category of topological groupoids. The subcategory of $\operatorname{Gpr} d(T o p)$, of the groupoids $\mathbb{G}$ with $G_{0}=\{*\}$ (the trivial topological space), is the category of topological groups, $\operatorname{Grp}(T o p)$.

Classically, the category of internal categories cannot itself be internal (Cat $\notin C a t)$.
Proposition 4.4. $\operatorname{cat}(\mathcal{N}) \in \operatorname{cat}(\mathcal{N})$. In other words, the category of small categories is a small category, in NF.

[^89]Proof. The functional relations and sets in the diagram defining $\operatorname{cat}(\mathcal{N})$ are homogeneous and definable as stratified set abstracts.

In fact, it would appear that all classically large categories of interest exist in $\operatorname{cat}(\mathcal{N})$, including $\mathcal{N}$ itself, when NF is taken as the base set theory for category theory. $\operatorname{cat}(\mathcal{N})$ is "closed," in terms of size. Nevertheless, there is a price for this - as can be seen in Lemma 4.5-relating to internal vs. external functors in NF.

In Section 3.1 we showed the distinction between internal and external functors in NF as the distinction between functors whose action maps are sets of NF and lateral functors in ML (see Definition 3.8).

Lemma 4.5. Define $\tilde{\mathcal{N}}$ as identical to $\mathcal{N}$, but viewed as a member of $\operatorname{cat}(\mathcal{N})$. The category of internal endofunctors $\operatorname{cat}(\mathcal{N})[\tilde{\mathcal{N}}, \tilde{\mathcal{N}}]$ is also an internal category, but is not (externally) equivalent to $[\mathcal{N}, \mathcal{N}]$.

Proof. If $T: \mathcal{N} \rightarrow \mathcal{N}$ has a graph in $\mathcal{N}$, one obtains the Burali-Forti paradox. Therefore, the inclusion of $\operatorname{cat}(\mathcal{N})[\tilde{\mathcal{N}}, \tilde{\mathcal{N}}]$ in $[\mathcal{N}, \mathcal{N}]$ is a proper embedding ${ }^{6}$

Remark ("Smallness" in $\mathcal{N}$ ). While the aforementioned distinction between internal and external in NF arises primarily as a distinction between functor categories, there are nonfunctor categories of importance that are not in $\operatorname{cat}(\mathcal{N})$. As mentioned earlier, in the case of NF, category theoretic smallness does not correspond to set theoretic smallness (i.e. it does not correspond to smallness in the sense of cardinality or closure under the powerset operation). Therefore, where we are tempted to refer to a category $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$ as "small" (as we would for $\mathbb{C} \in \operatorname{cat}($ Set $)$ ), we will simply refer to it as internal.

We use the phrase NF-small to denote a category that is small in the sense of NF - that is, in the sense that its collection of morphisms is a strongly cantorian set. $7^{7}$

Definition 4.6. An internal category $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$ is said to be $N F$-small if $C_{0}, C_{1} \in S C$. Thus, the category of NF-small categories is $\operatorname{cat}(S C) \subset \operatorname{cat}(\mathcal{N})$.

[^90]Similarly, a category $\mathcal{C}$ can be said to be locally NF-small if each of its hom-sets is strongly cantorian.

Example 4.7. As any hom-set between strongly cantorian sets is itself strongly cantorian, $S C$ itself provides an example of a locally NF-small (external) category.

## Internal Presheaves

Typically, as is the case where $\mathcal{E}=S$ et, neither $\operatorname{cat}(\mathcal{E})$ nor $\mathcal{E}$ are $\mathcal{E}$-internal categories. Nevertheless, one can give an alternative presentation of presheaves, internally. The closure of $\operatorname{cat}(\mathcal{N})$, in terms of size, extends to categories of internal presheaves. This gives $N F$ the unique feature: $\mathcal{N}^{\mathbb{C}} \in \operatorname{cat}(\mathcal{N})$.

Definition 4.8. Given an internal category $\mathbb{C} \in \operatorname{cat}(\mathcal{E})$, the category $\mathcal{E}^{\mathbb{C}}$ of (covariant) internal presheaves is defined as follows:

Objects: An internal presheaf $F=\left\langle F_{0}, e, \gamma\right\rangle$ is a commutative diagram:

subject to the following conditions:

$$
\begin{array}{lr}
e \circ\langle 1, i \circ \gamma\rangle=1_{F_{0}} & \left(F\left(1_{c}\right)=1_{F(c)}\right) \\
e \circ e \times 1=e \circ 1 \times m & (F(g f)=F(g) F(f))
\end{array}
$$

Morphisms: A map of internal functors $h: F \rightarrow G$ is a morphism $h: F_{0} \rightarrow G_{0}$ over $C_{0}$, which commutes with the respective action maps $e_{F}$ and $e_{G}$, satisfying naturality conditions.

In the case of Set, for a small category $\mathcal{C}$ and a presheaf $F: \mathcal{C} \rightarrow$ Set, we intuit the internal presentation of $F$ as the triple $\left\langle F_{0}, e, \gamma\right\rangle$ with the following correspondence:

- $F_{0}=\coprod_{c \in C_{0}} F(c)$
- For $x \in F(c), \gamma:\langle x, c\rangle \mapsto c$

Hence the object part of $F$ corresponds to $\gamma^{-1}$

- Given a morphism $f: c \rightarrow c^{\prime}$ in $\mathcal{C}$ and an object $x \in F(c),\langle f, x\rangle \in C_{1} \times_{d_{0}} F_{0}$. The image $F(f)(x) \in F\left(c^{\prime}\right)$ is defined by the action map $e: C_{1} \times_{d_{0}} F_{0} \rightarrow F_{0}$

The functoriality conditions imposed on $\left\langle F_{0}, e, \gamma\right\rangle$ turn out to be equivalent to algebra conditions for a monad induced by a free-forgetful adjunction. Under this correspondence, free algebras are equivalent to coproducts of representable presheaves.

Example 4.9 (Categories of Elements). Internal presheaves correspond to (external) functors, but also to a specific class of internal categories. Given an internal presheaf $\left\langle F_{0}, e, \gamma\right\rangle$, the following diagram, $\mathbb{F}$, is an object of $\operatorname{cat}(\mathcal{E})$ :

$$
C_{1} \times_{d_{0}} C_{1} \times_{d_{0}} F_{0} \xrightarrow{1 \times e} C_{1} \times_{d_{0}} F_{0} \underset{e}{\stackrel{\pi_{2}}{\leftrightarrows}\langle\gamma, 1\rangle \longrightarrow} F_{0}
$$

Internal categories arising in this way form a subcategory of $\operatorname{cat}(\mathcal{E})$, which we refer to as $\operatorname{Elts}(\mathbb{C})$.

There is an alternative characterisation of internal presheaves known as discrete opfibrations. That is, as a subcategory of $\operatorname{cat}(\mathcal{E}) / \mathbb{C}$ corresponding to internal functors, $\gamma: \mathbb{F} \rightarrow \mathbb{C}$, over $\mathbb{C}$, such that the following diagram is a pullback:


We can connect these two characterisations by the domain functor:

$$
\delta_{0}: \operatorname{cat}(\mathcal{E}) / \mathbb{C} \rightarrow \operatorname{cat}(\mathcal{E})
$$

Restricting $\delta_{0}$ to $\operatorname{OpFib}(\mathbb{C})$, we obtain an equivalence of categories:

$$
\delta_{0}: \operatorname{OpFib}(\mathbb{C}) \rightarrow E l t s(\mathbb{C})
$$

The internal presentation of (external) presheaves rests on the more basic equivalence between slice categories and mappings: $\mathcal{E} / X \cong[X, \mathcal{E}]$. In NF, the stratified equivalence
is: $\mathcal{N} / X \cong[T X, \mathcal{N}]$. This extends to a correspondence between the internal presheaf category $\mathcal{N}^{\mathbb{C}}$ - which itself a $\mathcal{N}$-internal category - and the internal functor category $\operatorname{cat}(\mathcal{N})[T \mathbb{C}, \tilde{\mathcal{N}}]$, where $T \mathbb{C}$ is induced by applying the $T$-functor to the diagram representing $\mathbb{C}$. But the increase in dimension (categorically speaking) adds some nuance to the result. Something we will make precise in Section 4.3.

## Basic Internal Colimits

The final aspect of elementary internal category theory, reviewed here, is the formation of (co)limits.$^{8}$ Given some internal category $\mathbb{C}$, a basic question one could ask is: does $\mathbb{C}$ have (co)limits of some small diagram?

In both $\operatorname{cat}(\mathcal{E})$ and $\mathcal{E}^{\mathbb{C}}$, finite limits are simply inherited from the ambient category (i.e. they are created by the obvious functors, which forget categorical structure). Colimits require the existence of reflexive coequalisers in $\mathcal{E}$ - reflexive, as $i$ splits both $d_{0}$ and $d_{1}$ in any internal category.

Definition 4.10. Given an internal category $\mathbb{C}, \underset{\longrightarrow}{\lim } \mathbb{C}$ is a coequalizer of $d_{0}$ and $d_{1}$, the domain and codomain maps of $\mathbb{C}$.

$$
C_{1} \xrightarrow[d_{1}]{\stackrel{d_{0}}{\longrightarrow}} C_{0}-->\underset{\longrightarrow}{\lim } \mathbb{C}
$$

$\xrightarrow{\lim }: \operatorname{cat}(\mathcal{E}) \rightarrow \mathcal{E}$ is left adjoint to the inclusion of discrete internal categories $i: \mathcal{E} \hookrightarrow$ $\operatorname{cat}(\mathcal{E})$.

We define the colimit of an internal presheaf $F \in \mathcal{E}^{\mathbb{C}}$ by applying $\underset{\longrightarrow}{\lim }$ to the internal category of elements, $\mathbb{F}$. We can then define a functor:

$$
\underset{\mathbb{C}}{\lim }: \mathcal{E}^{\mathbb{C}} \rightarrow \mathcal{E}
$$

which, as in the external case, is left adjoint to the constant presheaf functor defined by the action $X \mapsto\left\langle X \times C_{0}, 1 \times d_{1}, \pi_{2}\right\rangle$. The existence of $\underline{l i m}_{\mathbb{C}}$, for any object in $\operatorname{cat}(\mathcal{E})$,

[^91]is referred to as internal cocompleteness. A diagram is small if it is the image of an internal category (i.e. an internal presheaf).

Existence of reflexive coequalizers in $\mathcal{N}$, as with coequalizers in general, appears to require assumptions beyond the axioms of NF. Thus, we are not able to prove that $\mathcal{N}$ is internally cocomplete $?^{9}$ We can prove a series of partial results, such as:

Proposition 4.11. Given an internal category $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$ with an object $\pi \mathbb{C}$ of path connected components in the image of $T$ (up to isomorphism), the coequalizer $\xrightarrow{\lim } \mathbb{C}$ exists. Thus, $\mathcal{N}$ has colimits for any diagram $F$, where $\pi \mathbb{F}$ is in the (essential) image of $T$.

### 4.2 Standard Internal Presentation of Presheaves

## The Yoneda Lemma

The eponymous embedding and lemma of Yoneda is, effectively, a generalization of Cayley's Theorem from Group Theory to Category Theory ${ }^{10}$ Such a statement, ubiquitous in textbooks and Part III maths lectures, is not unlike the lemma it is about: intriguing and unapologetically brief. That said, there are many helpful accounts [4, 31]. We begin with a summary of some main points, tailored to NF.

Definition 4.12. The Yoneda Embedding $\mathcal{Y}_{\mathcal{C}}: \mathcal{C}^{o p} \rightarrow S e t^{\mathcal{L}}$ is defined:

$$
\begin{array}{lr}
C \mapsto \mathcal{C}(C,-): \mathcal{C} \rightarrow \text { Set } & \text { (Action on Objects) } \\
\left(f: C \rightarrow C^{\prime}\right) \mapsto-\circ f: \mathcal{C}\left(C^{\prime},-\right) \rightarrow \mathcal{C}(C,-) & \text { (Action on Morphisms) }
\end{array}
$$

Lemma 4.13 (Yoneda's Lemma). Given a (locally) small category $\mathcal{C}$ and a covariant presheaf $F: \mathcal{C} \rightarrow$ Set, there is a natural isomorphism:

$$
\operatorname{Nat}(\mathcal{C}(C,-), F) \cong F(C)
$$

[^92]Proof. (Sketched) Consider an arbitrary natural transformation $\omega: \mathcal{C}(C,-) \rightarrow F$ and a morphism $f: C \rightarrow C^{\prime}$.


Consider the element $i d_{C} \in \mathcal{C}(C, C)$, the commutativity of the above diagram implies that $\omega_{C^{\prime}}(f)=F(f)\left(\omega_{C}\left(i d_{C}\right)\right)$. Thus the action of $\omega$ is determined uniquely by a single choice: where $\omega_{C}$ maps the identity morphism. Thus, there are as many natural transformations as there are choices (i.e. elements of $F C$ ).

As an immediate corollary, $\mathcal{Y}$ is an embedding.
Remark (Typing Considerations). From the standpoint of NF, $\mathcal{Y}$ is an external (stratified, but type-raising) functor. The natural isomorphism is immediately problematic. A natural transformation $\omega$ is a set of maps, each of which would have the same type as $F C$ in a stratification, thus we are attempting to assert an isomorphism between objects that are 2 "types" apart. At the same time, however, there is something tempting about the Yoneda Lemma: it seems like a tool that we could use to "type-lower" in the appropriate context. Further, $\mathcal{Y}_{\mathcal{C}}$ embeds $\mathcal{C}^{o p}$ into $S e t^{\mathcal{C}}$, much in the way that $\{\cdot\}_{X}$ embeds $X$ into $P X$. We are led to consider the following claim: Yoneda is a T-functor. But such a statement requires further motivation of Yoneda itself.

We consider three successive generalizations ${ }^{[1]}$ Sets, Posets, Categories. We can think of enrichment (i.e. the hom-sets) in a similar succession: trivial (equality), $\mathbf{2}$ (the two element lattice), and Set. Turning now to just posets and categories, we see that the powerset operation has a natural generalization to $D \mapsto \downarrow D$ (the set of downward closed sets) and $S e t^{\mathcal{C o p}^{\circ}}$.

In the case of posets $D$ and $D^{\prime}$, consider a map $f: D \rightarrow \downarrow D^{\prime}$. This map has a unique

[^93]cocontinuous ${ }^{12}$ extension $\hat{f}: \downarrow D \rightarrow \downarrow D^{\prime}$, where $\hat{f} \circ\{\cdot\}=f$ completes the diagram.


The map $\{\cdot\}$ is the map $\downarrow d$ for individual elements of $D$, and we can consider the embedding $\{\cdot\}_{D}: D \rightarrow \downarrow D$ the cocompletion of $D$.

For categories $\mathcal{C}$ and $\mathcal{C}^{\prime}$, we obtain the same extension of a functor $F: \mathcal{C} \rightarrow S e t^{\mathcal{C}^{\circ p^{\prime}}}$ to a cocontinuous (colimit preserving) $L_{\mathcal{Y}_{\mathcal{C}}} F: S e t^{\mathcal{C O D}^{o p}} \rightarrow S e t^{\mathcal{C o p}^{o p}}$, where $C \mapsto \mathcal{C}(-, C)$ is the embedding of $\mathcal{C}$ in its cocompletion Set ${ }^{\mathcal{C}^{\text {op }}}$. Thus, we can view the Yoneda Lemma, in full generality, as the category theoretic analogue to $T: X \rightarrow P X$.

### 4.2.1 The Internal Yoneda Lemma

Definition 4.14. For an internal category $\mathbb{C}, U: \mathcal{E}^{\mathbb{C}} \rightarrow \mathcal{E} / C_{0}$ is the forgetful functor, which takes an internal presheaf $\left(F_{0}, \gamma_{0}, e\right)$ to its object component $\gamma_{0}: F_{0} \rightarrow C_{0}$.

Definition 4.15. For an internal category $\mathbb{C}, R: \mathcal{E} / C_{0} \rightarrow \mathcal{E}^{\mathbb{C}}$ takes an object $(X, \gamma)$ of $\mathcal{E} / C_{0}$ to the (representable) internal presheaf $R(\gamma)=\left(X \times_{d_{0}} C_{1}, d_{1} \circ \pi_{2}, 1 \times m\right)$, where $X \times{ }_{d_{0}} C_{1}$ is the pullback of $(X, \gamma)$ along $d_{0}$.


One can easily check that $R(\gamma)$ satisfies the conditions of a discrete opfibration, as defined in the prior section, with $1 \times m$ being the appropriate action map. ' $R$ ' stands for representable, a claim easily justified if one considers $R\left(c: 1 \rightarrow C_{0}\right)$ for a global element of $C_{0}$. If $\mathbb{C}$ is a small category in the classical sense, $d_{1} \circ \pi_{2}: 1 \times_{d_{0}} C_{1} \rightarrow C_{0}$ has as a fibre over each $c^{\prime} \in C_{0}$, the collection:

$$
\left\{f \mid f \in C_{1} \wedge d_{0}(f)=c \wedge d_{1}(f)=c^{\prime}\right\}
$$

[^94]which is clearly $\mathcal{C}\left(c, c^{\prime}\right)$. The action map $1 \times m$ just defines post-composition. Thus $R(c)$ is precisely the display of $\mathcal{C}(C,-)$. In general, $\gamma: X \rightarrow C_{0}, R(\gamma)$ is just a $\gamma$-indexed coproduct of representable functors.

Lemma 4.16 (Free-Forgetful: Presheaves). [22] For an internal category $\mathbb{C}, R \dashv U$.

Proof. Straightforward verification. For later use, we note the unit of this adjunction:

$$
\eta_{\gamma}:(X, \gamma) \rightarrow\left(X \times_{d_{0}} C_{1}, d_{1} \circ \pi_{2}\right)
$$

is equivalent to the map $\langle 1, i \circ \gamma\rangle$.

Corollary 4.17. The bijection between hom-sets:

$$
\mathcal{E}^{\mathbb{C}}\left(R(\gamma),\left(F_{0}, \tau, e\right)\right) \cong \mathcal{E} / C_{0}(\gamma, \tau)
$$

induces the Yoneda Lemma.

Proof. Consider the bijection, where $\gamma=c: 1 \rightarrow C_{0}$ :

$$
\mathcal{E}^{\mathbb{C}}\left(R(c),\left(F_{0}, \tau, e\right)\right) \cong \mathcal{E} / C_{0}(c, \tau)
$$

### 4.2.2 Presheaves as Algebras

The adjunction $R \dashv U$ defines representable presheaves as free structures. Presheaves, in general, are defined as the algebras to the canonical monad induced by the adjunction. In other words, (internal) presheaves are quotients of representable presheaves.

Theorem 4.18. [22] Given a finitely complete category $\mathcal{X}$ and $\mathbb{C} \in \operatorname{cat}(\mathcal{X}), U: \mathcal{X}^{\mathbb{C}} \rightarrow$ $\mathcal{X} / C_{0}$ is monadic.

Proof. Sketched. The monad is just the monad induced by $R \dashv U$, but it is convenient to observe that $T_{\mathbb{C}}=\Sigma_{d_{1}} d_{0}^{*}: \mathcal{X} / C_{0} \rightarrow \mathcal{X} / C_{0}$. The unit component at some $\gamma$ is given
by: $\eta_{\gamma}^{T}=\langle 1, i \gamma\rangle . \mu_{\gamma}^{T}: T_{\mathbb{C}}^{2}(\gamma) \rightarrow T_{\mathbb{C}}(\gamma)$ is given by $1 \times m$, where $m$ is the composition map for $\mathbb{C}$. Notice that $(1 \times m) \circ\langle 1, i \gamma\rangle=i d_{\gamma}$ is a property of both monads and of internal presheaves $\left(F(i d)=i d_{F}\right)$.

Given an internal presheaf $\left(F_{0}, \gamma_{0}, e\right)$, the action map $e: F_{0} \times_{d_{0}} C_{1} \rightarrow F_{0}$ satisfies $e(e \times 1)=e(1 \times m)$ and $e \times 1=T_{\mathbb{C}}(e)$. Thus, functoriality conditions are algebra conditions, yielding the split coequalizer:

$$
C_{1} \times_{d_{0}} C_{1} \times_{d_{0}} F_{0} \underset{\left\langle 1, i \circ d_{1} \circ \pi_{2}\right\rangle}{\stackrel{e \times 1}{\Psi} \longrightarrow} C_{1} \times_{d_{0}} F_{0} \underset{\langle 1, i \gamma\rangle=\eta_{\gamma}^{T}}{\rightleftarrows} F_{0}
$$

$e \times 1$ and $1 \times m$ are algebra maps between free algebras. More than that, they are truly "free" in the sense that $e$ being an algebra map has nothing to do with $e \times 1$ being a map of free algebras. Naturality of $\mu^{T}$ is sufficient to guarantee this.

Remark (Presheaves as Algebras in NF). Quotients are known to be problematic in any stratified theory. But, in the case of $R(X, \gamma)$, the coequalizer defining the quotient is split by $\eta_{R(\gamma)}$. From the perspective of stratification, the splitting gives a canonical selection function, allowing one to form the coequalizer diagram ${ }^{[13}$ Thus, $\mathcal{N}$ has "enough" quotients to form (internal) presheaves as $T_{\mathbb{C}}$-algebras ${ }^{[14}$

### 4.2.3 Presheaves as Coalgebras

In a topos, the presheaf monad $T_{\mathbb{C}}$ has a right adjoint $S_{\mathbb{C}}=\Pi_{d_{0}} d_{1}^{*}$, induced by composition of adjunctions $\Sigma_{d_{1}} \dashv d_{1}{ }^{*}$ and $d_{0}{ }^{*} \dashv \Pi_{d_{0}}$. Therefore, the category of $T_{\mathbb{C}}$-algebras is equivalent to the category of $S_{\mathbb{C}}$-coalgebras ${ }^{15}$ In other words, presheaves have equivalent presentations as algebras and coalgebras.

[^95]The following isomorphism, describing how sums and products commute with homsets, provides some intuition for the equivalence between the monadic and comonadic presentations:

$$
\mathcal{C}\left(\coprod_{J} a_{j}, c\right) \cong \prod_{J} \mathcal{C}\left(a_{j}, c\right)
$$

For an arbitrary (internal) presheaf $F=\left(F_{0}, \gamma_{0}, e\right)$, we give the explicit algebraic and coalgebraic presentations $\stackrel{10}{16}^{16}$

The monadic description, $T_{\mathbb{C}}=\Sigma_{d_{1}} d_{0}{ }^{*}$ :

$$
e_{C}: \coprod_{d_{1}(f)=C} F\left(d_{0}(f)\right) \rightarrow F(C) ; x_{f} \mapsto F(f)(x)
$$

The comonadic description, $S_{\mathbb{C}}=\Pi_{d_{0}} d_{1}{ }^{*}$ :

$$
l_{C}: F C \rightarrow \prod_{d_{0}(g)=C} F\left(d_{1}(g)\right) ; x \mapsto(g \mapsto F(g)(x))
$$

The $T_{\mathbb{C}}$-algebra conditions for $e$ and $S_{\mathbb{C}}$-coalgebra conditions for $l$ are exactly the functoriality conditions for $F$.

### 4.3 The Internal Presentation of Presheaves in $\mathcal{N}$

### 4.3.1 The NF Yoneda Lemma

Theorem 4.19 (NF Yoneda). Given an NF-small category $\mathcal{C}$ and a covariant presheaf $F: \mathcal{C} \rightarrow \mathcal{N}:$

$$
\operatorname{Nat}(\mathcal{C}(C,-), F) \cong T F(C)
$$

forms an isomorphism, natural in $C$ and $F$.

Proof. Consider the global element $c: 1 \rightarrow C_{0}$ corresponding to the object $C$ as above. The adjunction $R \dashv U$ above yields an (internal) isomorphism:

$$
\mathcal{N}^{\mathbb{C}}(R(c), F) \cong \mathcal{N} / C_{0}(c, U(F))
$$

[^96]$R(c)$ is precisely $\mathcal{C}(c,-)$, as described earlier. $\mathcal{N} / C_{0}(c, U(F))$ is the set of morphisms over $C_{0}$ between $(1, c)$ and $\left(F_{0}, \gamma_{0}\right)$, where the latter is the object component of $F$. Thus, we obtain precisely the set of global elements of $F_{0}$ in the fibre of $\gamma_{0}$ over $C$. This is equivalent to the set $\mathcal{N}(1, F(C))$, where $\mathrm{F}(\mathrm{C})$ is defined explicitly in the pullback diagram below. We conclude:
$$
\mathcal{N}^{\mathbb{C}}(R(c), F) \cong \mathcal{N} / C_{0}(1, F(C)) \cong T F(C)
$$


In the case where $F$ is also induced by a global element $c^{\prime}: 1 \rightarrow C_{0}$ (i.e. $F=R\left(c^{\prime}\right)$ ), the associated embedding is full and faithful.

## (Relative) Algebraic Presheaves in $\mathcal{N}$

$\mathcal{N}$, being finitely complete, has algebraic presentations of internal presheaves. The existence of a comonadic presentation is far less clear as we have shown $\mathcal{N}$ does not permit standard dependent products. However, given the intuitive duality, combined with the role of comonadic presentations in proving internal presheaf categories (of toposes) are toposes, we see a need to develop a form in $\mathcal{N} .{ }^{17}$

The natural generalization of $S_{\mathbb{C}}$ to $\mathcal{N}$ arises from the composition of dependent sums with modified-dependent products. To form the appropriate relative comonad $\tilde{S}_{\mathbb{C}}$, we use the more general results of Chapter 2.

### 4.3.2 The Presheaf Relative Comonad

$$
\begin{align*}
& \left(\Sigma_{d_{1}} \dashv d_{1}{ }^{*} \wedge d_{0}{ }^{*} \dashv \Pi_{d_{0}}\right) \Longrightarrow\left(T_{\mathbb{C}}=\Sigma_{d_{1}} d_{0}{ }^{*}\right) \dashv\left(\Pi_{d_{0}} d_{1}{ }^{*}=S_{\mathbb{C}}\right)  \tag{Topos}\\
& \left(\Sigma_{d_{1}} \dashv d_{1}{ }^{*} \wedge d_{0}{ }^{*}{ }_{T} \dashv \tilde{\Pi}_{d_{0}}\right) \Longrightarrow\left(T_{\mathbb{C}}=\Sigma_{d_{1}} d_{0}{ }^{*}\right)_{T} \dashv\left(\tilde{\Pi}_{d_{0}} d_{1}{ }^{*}=\tilde{S}_{\mathbb{C}}\right) \tag{NF}
\end{align*}
$$

[^97]Remark (Deriving Coalgebraic Presentations from Symmetric Lifts). The following results are corollary to Lemma 2.10 and 2.11 .

Corollary 4.20. In $\mathcal{N}, T_{\mathbb{C}}$ is a relative $T$-left adjoint:

$$
\Sigma_{d_{1}} d_{0}{ }^{*}{ }_{T} \dashv \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}
$$

for any $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$.
Corollary 4.21. Given any $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, there is a relative $T$-right adjunction:

$$
T d_{0}{ }^{*} \Sigma_{T d_{1}} \dashv_{T} T d_{1}{ }^{*} \tilde{\Pi}_{d_{0}}
$$

Remark (The Need for Symmetric Lifts). Following the classical case, one would expect that (one of) the above results would be sufficient to form the (relative) coalgebraic presentation. But, in this case, a relative adjunction alone does not preserve a sufficient level of adjoint symmetry.

In order to form a relative comonad, one requires a $T$-relative right adjoint to $\Sigma_{T d_{1}} T d_{0}{ }^{*}$. This turns out to be the completion of a symmetric lift, whose upper relative adjoint corresponds to Corollary 4.20. Thus, we obtain a further piece of empirical evidence for the advantages of working with symmetric lifts, as a generalization of adjoint symmetry: The relative coalgebraic presentation of presheaves arises not as a composite of relative adjoints, but as a composite of symmetric lifts.

$$
\begin{aligned}
& \mathcal{N} / C_{0} \xrightarrow{d_{0}{ }^{*}} \mathcal{N} / C_{0} \stackrel{\Sigma_{d_{1}}}{\stackrel{d_{1}{ }^{*}}{ }} \mathcal{N} / C_{0} \quad \mathcal{N} / C_{0} \xrightarrow{\Sigma_{d_{1} d_{0}{ }^{*}}^{\longrightarrow}} \mathcal{N} / C_{0} \\
& T_{C_{0}} \downarrow \quad \check{\Pi}_{d_{0}}{ }^{T_{C_{0}}} \downarrow \stackrel{T \circ d_{1}{ }^{*}=T d_{1}^{*} \circ T}{{ }^{*}} \downarrow^{T_{C_{0}}}=T_{C_{0}} \downarrow \quad \tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \quad \downarrow^{T_{C_{0}}}
\end{aligned}
$$

Theorem 4.22. Given a category $\mathbb{C} \in \operatorname{Cat}(\mathcal{N})$, there is a $T$-relative right adjunction $\Sigma_{T d_{1}} T d_{0}{ }^{*} \dashv_{T} \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}$, which completes the symmetric lift:

We note the following lemma, before proving Theorem 4.24.

Lemma 4.23. In $\mathcal{N}, T$ commutes with modified-dependent products. Thus, the following diagram commutes for any $f: C \rightarrow D$ :


Proof. This follows from the construction of dependent products. It suffices to observe that $\tilde{\Pi}_{T f} \circ T(\gamma)$ consists of sections $g:(T f)^{-1}(\{c\}) \rightarrow T A$. As $(T f)^{-1}(\{c\})=T\left(f^{-1}(c)\right)$ and $T$ is full and faithful, $g=T h$ for some $h$ in $\tilde{\Pi}_{f}(\gamma)$.

The following result could be stated as a corollary to Theorem 2.33 and Theorem 4.22, the general relative comonadicity theorem. However, in an effort to better understand relative coalgebraic presentations of presheaves in NF, we carry out the proof set theoretically.

Theorem 4.24 (The Presheaf Relative Comonad, $\tilde{S}_{\mathbb{C}}$ ). For any $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, there is a relative comonad $\tilde{S}_{\mathbb{C}}=\left(\tilde{\Pi}_{d_{0}} d_{1}{ }^{*}, \bar{\varepsilon},(\hat{)})\right.$, corresponding to the coalgebraic presentation of internal presheaves over $\mathbb{C}$.

Proof. (Set Theoretic) The proof is carried out in three parts.

1. Define the Relative co-unit, $\bar{\varepsilon}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \Rightarrow T$
2. Define Relative Comultiplication: $\hat{( }):\left|\tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \downarrow T\right| \rightarrow\left|\tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \downarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}\right|$
3. Prove the Relative Comonad Conditions.

## 1) The Relative Co-unit

Using set theoretic notation, each component $\bar{\varepsilon}_{c}: \prod_{d_{0}(g)=c} F\left(d_{1}(g)\right) \rightarrow T F(c)$ is defined by the action $\langle f,\{c\}\rangle \mapsto\left\{\pi_{2}\left(f\left(i d_{c}\right)\right)\right\}$, where $f$ is a section of $\pi_{1}: C_{1} \times_{d_{1}} \coprod_{c} F(c) \rightarrow C_{1}$,
over $d_{0}{ }^{-1}(c)$. In other words, each element of the indexed product is mapped to the singleton set containing its component at the identity morphism of $c$.

Using category theoretic notation and the presentation of indexed families as fibres, we obtain the morphism in $\mathcal{N} / T C_{0}$ :

where:

$$
\tilde{\Pi}_{d_{0}} d_{1}^{*}(\gamma)=\left\{\langle g,\{c\}\rangle \mid g: d_{0}^{-1}(c) \rightarrow C_{1} \times_{d_{1}} D \wedge \pi_{1} \circ g=i d_{C_{1}} \upharpoonright d_{0}^{-1}(c)\right\}
$$

and:

$$
\bar{\varepsilon}_{\gamma}: \tilde{\Pi}_{d_{0}} d_{1}^{*}(\gamma) \rightarrow T D ;\langle g,\{c\}\rangle \mapsto\left\{\pi_{2} \circ g\left(i d_{c}\right)\right\}
$$

## 2a) Comultiplication

In the classical case (say, in the set theory of $\mathrm{ZF}(\mathrm{C})$ ), we form the comultiplication map by its components:

$$
\delta_{c}: \prod_{d_{0}(f)=c} F\left(d_{1}(f)\right) \rightarrow \prod_{d_{0}(f)=c} \prod_{d_{1}(f)=d_{0}(g)} F\left(d_{1}(g)\right)
$$

Each element of the first dependent product can be viewed as a map:

$$
h: d_{0}^{-1}(c) \rightarrow \coprod_{d_{0}(f)=c} F\left(d_{1}(f)\right)
$$

which assigns to each $j \in d_{0}{ }^{-1}(c)$ an element $h(j) \in F\left(d_{1}(j)\right)$. $\delta_{c}(h)$ is the map:

$$
\delta_{c}(h): d_{0}^{-1}(c) \rightarrow \coprod_{d_{0}(f)=c}\left[d_{0}^{-1} \circ d_{1}(f) \Longrightarrow \coprod_{d_{0}(g)=d_{1}(f)} F\left(d_{1}(g)\right)\right]
$$

which maps each $f \in d_{0}^{-1}(c)$ to an element $h_{f} \in \prod_{d_{1}(f)=d_{0}(g)} F\left(d_{1}(g)\right)$ :

$$
\delta_{c}(h):\left(f: c \rightarrow c^{\prime}\right) \mapsto h_{f}: d_{0}^{-1}\left(c^{\prime}\right) \rightarrow \coprod_{d_{0}(g)=c^{\prime}} F\left(d_{1}(g)\right)
$$

defined by the action:

$$
\delta_{c}(h)(f) \equiv h_{f}:\left(g: c^{\prime} \rightarrow c^{\prime \prime}\right) \mapsto h(g \circ f) \in F\left(c^{\prime \prime}\right)
$$

Thus, for any morphism of $\mathbb{C}, f: c \rightarrow c^{\prime}, h$ is mapped to a function $h_{f}$, whose action is determined by that of $h$, on the subset of the fibre $d_{0}^{-1}(c)$, determined by the members of $d_{0}^{-1}\left(c^{\prime}\right)$, precomposed with $f$. Thus, the comultiplication condition of the comonad expresses that functors (hence presheaves) preserve composition.

## 2b) Relative Comultiplication

In NF, the components of $\delta$ would be unstratified. As discussed in Chapter 2, this is a special case of a more general issue. One is attempting to iterate a functor which is not necessarily an endofunctor ${ }^{18}$

The relative-comultiplication map. ${ }^{19}$

$$
\hat{()}:\left|\tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \downarrow T\right| \rightarrow\left|\tilde{\Pi}_{d_{0}} d_{1}{ }^{*} \downarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}\right|
$$

returns $\hat{k}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta)$ for any map $k: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow T \beta$, in $\mathcal{N} / T C_{0}$ :

with $\gamma: X \rightarrow C_{0}$ and $\beta: B \rightarrow C_{0}$.

Elements of $\tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma)$ are sections of $\pi_{2}$, over fibres of $d_{0}$. Given $\langle g,\{c\}\rangle \in \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma)$, any morphism $j: c \rightarrow c^{\prime}$ in $\mathbb{C}$ determines another element $\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle \in \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma)$ :

$$
g_{j} \equiv\left\langle i d_{C_{1}}, \pi_{2} \circ g(-\circ j)\right\rangle: d_{0}^{-1}\left(c^{\prime}\right) \rightarrow C_{1} \times_{d_{1}} X
$$

As $\hat{k}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta)$ is a map over $T C_{0}, \pi_{2} \circ \hat{k}=\pi_{2}$. Thus, $\hat{k}(\langle g,\{c\}\rangle)$ is defined by the following equality and actions:

$$
\begin{array}{lr}
\pi_{2} \circ \hat{k}(\langle g,\{c\}\rangle)=\{c\} & \left(\hat{k} \text { is a map over } T C_{0}, \text { so } \pi_{2} \circ \hat{k}=\pi_{2}\right) \\
\pi_{1} \circ \hat{k}(\langle g,\{c\}\rangle): d_{0}{ }^{-1}(c) \rightarrow C_{1} \times_{d_{1}} B & \left(\langle g,\{c\}\rangle \in \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma)\right) \\
\pi_{1} \circ \hat{k}(\langle g,\{c\}\rangle):\left(j: c \rightarrow c^{\prime}\right) \mapsto\left\langle j, \cup k\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right\rangle & \text { (action of } \left.\pi_{1} \circ \hat{k}(\langle g,\{c\}\rangle)\right) \\
g_{j}:\left(w: c^{\prime} \rightarrow c^{\prime \prime}\right) \mapsto\left\langle w, \pi_{2} g(w \circ j)\right\rangle & \text { (action of } \left.g_{j}\right)
\end{array}
$$

[^98]As $k: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow T \beta$ is a morphism over $T C_{0}$ :

$$
T \beta \circ k\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)=\left\{c^{\prime}\right\}
$$

Therefore, the map $\pi_{1} \circ \hat{k}(\langle g,\{c\}\rangle)$ is a section of $\pi_{1}: C_{1} \times_{d_{1}} B \rightarrow C_{1}$ over $d_{0}{ }^{-1}\left(c^{\prime}\right)$. Thus, $\hat{k}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta)$ is well-defined as a map over $T C_{0} .{ }^{20}$

## 3) The Three Conditions Defining a Relative Comonad

Note: As $\hat{k}$ is a map in $\mathcal{N} / T C_{0}$, it is always the case that $\pi_{2} \circ \hat{k}=\pi_{2}$. Thus, we adopt the convention:

$$
\hat{k}(\langle g,\{c\}\rangle) \equiv\langle\hat{k}(g),\{c\}\rangle
$$

(3.1) The first condition requires that, for any $\beta$ in $\mathcal{N} / C_{0}, \hat{\bar{\varepsilon}}_{\beta}=i d_{\tilde{\Pi}_{0} d_{1}{ }^{*}(\beta)} .^{21}$

Given an element $g \in \tilde{\Pi}_{0} d_{1}{ }^{*}(\beta)$,

$$
\bar{\varepsilon}\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)=\left\{\pi_{2} \circ g_{j}\left(i d_{c^{\prime}}\right)\right\}
$$

Hence:

$$
\cup \bar{\varepsilon}\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)=\pi_{2} g\left(i d_{c^{\prime}} \circ j\right)=\pi_{2} g(j)
$$

$\hat{\varepsilon}(g)=g$ is then obtained from the following chain of equalities:

$$
\hat{\bar{\varepsilon}}(g)(j)=\left\langle j, \cup \bar{\varepsilon}\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right\rangle=\left\langle j, \pi_{2}(g(j))\right\rangle=g(j)
$$

(3.2) $\bar{\varepsilon}_{\beta} \circ \hat{k}=k$ requires the following diagram to commute over $T C_{0}$ :


[^99]We prove this by the following chain of equivalences:

$$
\begin{array}{lr}
\bar{\varepsilon}_{\beta} \circ \hat{k}(\langle g,\{c\}\rangle)=\left\{\pi_{2} \circ \hat{k}(g)\left(i d_{c}\right)\right\} & \text { (definition of } \bar{\varepsilon}) \\
=\left\{\cup k\left(\left\langle g_{i d_{c}},\{c\}\right\rangle\right)\right\}=\{\cup k(\langle g,\{c\}\rangle)\} & \text { (definition of } \left.\hat{k}, g_{i d_{c}}=g\right) \\
=k(\langle g,\{c\}\rangle) & (T \circ \cup \circ T=T)
\end{array}
$$

(3.3) The co-distributive law:

$$
\hat{h} \circ \hat{k}=\widehat{h \circ \hat{k}}
$$

corresponds to the following commutative diagram, for any pair of morphisms $k$ : $\tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow T(\beta)$ and $h: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta) \rightarrow T(\alpha)$, in $\mathcal{N} / T C_{0}$.

where $\gamma: X \rightarrow C_{0}, \alpha: A \rightarrow C_{0}$ and $\beta: B \rightarrow C_{0}$.
$\hat{k}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta)$ is defined as above: ${ }^{22}$

$$
\left(j: c \rightarrow c^{\prime}\right) \longmapsto \stackrel{\hat{k}(g)}{\longrightarrow}\left\langle j, \cup k\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right\rangle
$$

All we can say, in general, about $h \circ \hat{k}$ is that $T(\alpha) \circ h \circ \hat{k}=\pi_{2}$ (it is a map over $T C_{0}$ ):

From $h \circ \hat{k}$, we obtain $\widehat{h \circ \hat{k}}: \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma) \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\alpha)$ :

$$
\left(j: c \rightarrow c^{\prime}\right) \stackrel{\widehat{h \circ \hat{k}}(g)}{\longrightarrow}\left\langle j, \cup h \circ \hat{k}\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right\rangle
$$

On the other hand, the composite $\hat{h} \circ \hat{k}$ is defined:

$$
\hat{h} \circ \hat{k}(\langle g,\{c\}\rangle)=\hat{h}(\langle\hat{k}(g),\{c\}\rangle)=\langle\hat{h}(\hat{k}(g)),\{c\}\rangle
$$

[^100]where the action of $\hat{h} \circ \hat{k}$ is defined:
$$
\left(j: c \rightarrow c^{\prime}\right) \stackrel{\hat{h}(\hat{k}(g))}{\longrightarrow}\left\langle j, \cup h\left(\left\langle\hat{k}(g)_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right\rangle
$$

The desired general equivalence:

$$
\hat{h} \circ \hat{k}=\widehat{h \circ \hat{k}}
$$

can now be reduced to:

$$
\left\langle\hat{k}(g)_{j},\left\{c^{\prime}\right\}\right\rangle=\hat{k}\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right) \equiv\left\langle\hat{k}\left(g_{j}\right),\left\{c^{\prime}\right\}\right\rangle
$$

Thus, we wish to prove $\hat{k}(g)_{j}=\hat{k}\left(g_{j}\right)$. First, recall the definition of $g_{j}$, where $j: c \rightarrow c^{\prime}$, by the action:

$$
\left(w: c^{\prime} \rightarrow c^{\prime \prime}\right) \stackrel{g_{j}}{\longrightarrow}\left\langle w, \pi_{2} \circ g(w \circ j)\right\rangle
$$

As $\langle g,\{c\}\rangle \in \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\gamma), \gamma \circ \pi_{2} \circ g(w \circ j)=c^{\prime \prime}$. Now consider the respective actions of $\hat{k}(g)_{j}$ and $\hat{k}\left(g_{j}\right)$ on $w: c^{\prime} \rightarrow c^{\prime \prime}:$

$$
\begin{aligned}
& \left(w: c^{\prime} \rightarrow c^{\prime \prime}\right) \longmapsto \begin{array}{|c}
\hat{k}\left(g_{j}\right)
\end{array}\left\langle w, \cup k\left(\left\langle\left(g_{j}\right)_{w},\left\{c^{\prime \prime}\right\}\right\rangle\right)\right\rangle \\
& \left(w: c^{\prime} \rightarrow c^{\prime \prime}\right) \longmapsto \hat{k}(g)_{j}
\end{aligned}\left\langle w, \pi_{2} \circ \hat{k}(g)(w \circ j)\right\rangle, ~ l
$$

For $\hat{k}(g)_{j}(w)$, note the equivalence (by definition of $\hat{k}$ ):

$$
\hat{k}(g)(w \circ j)=\left\langle w \circ j, \cup k\left(\left\langle g_{w \circ j},\left\{c^{\prime \prime}\right\}\right\rangle\right)\right\rangle
$$

Thus, we obtain the conditional:

$$
\left[\forall w \cdot g_{w \circ j}=\left(g_{j}\right)_{w}\right] \Longrightarrow \hat{k}(g)_{j}=\hat{k}\left(g_{j}\right)
$$

To show $g_{w \circ j}=\left(g_{j}\right)_{w}$, consider the following chain of equivalences, for an arbitrary map $t: c^{\prime \prime} \rightarrow c^{\prime \prime \prime}:$

$$
\begin{aligned}
& \left(g_{j}\right)_{w}(t)=\left\langle t, \pi_{2} g_{j}(t \circ w)\right\rangle \\
& =\left\langle t, \pi_{2}\left(\left\langle t \circ w, \pi_{2}(g(t \circ w \circ j))\right\rangle\right)\right\rangle \\
& =\left\langle t, \pi_{2}(g(t \circ w \circ j))\right\rangle \\
& =g_{w \circ j}(t)
\end{aligned}
$$

Definition (Relative $\tilde{S}_{\mathbb{C}}$-Coalgebras). For a given $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, a relative $\tilde{S}_{\mathbb{C}}$-coalgebra is a pair $\left(\gamma_{0}, \chi\right)$, where $\gamma_{0}: F_{0} \rightarrow T C_{0}$ and $\chi$ is an object function:

$$
\chi:\left|\gamma_{0} \downarrow T_{C_{0}}\right| \rightarrow\left|\gamma_{0} \downarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}\right|
$$

Satisfying the unit and distributive laws:


### 4.3.3 Algebraic and Relative Coalgebraic Presentations

In the classical case, the following are equivalent presentations of an internal presheaf $\left(F_{0}, \gamma_{0}, e\right)$ over $\mathbb{C} \in \operatorname{cat}(\mathcal{E})$.

1. A $\mathbf{T}_{\mathbb{C}}$-algebra: $e: \Sigma_{d_{1}} d_{0}{ }^{*}\left(\gamma_{0}\right) \rightarrow \gamma_{0}$
2. A $\mathbf{S}_{\mathbb{C}}$-coalgebra: $\tilde{e}: \gamma_{0} \rightarrow \Pi_{d_{0}} d_{1}{ }^{*}\left(\gamma_{0}\right)$
3. A Manes-style $\mathbf{S}_{\mathbb{C}}$-coalgebra: $\left(\gamma_{0}, \chi\right)$ where $\chi(f)$ is defined $\Pi_{d_{0}} d_{1}{ }^{*}(f) \circ \tilde{e}$

In $\mathcal{N}$, the second object is unstratified. The first object and the relative version of the third are stratified, but their relationship is not entirely clear.

The functor $\tilde{\Pi}_{d_{0}} d_{1}^{*}$, corresponding to the relative comonad $\tilde{S}_{\mathbb{C}}$, arises as a relative right adjoint to $\Sigma_{T d_{1}} T d_{0}{ }^{*}$. Therefore, the partial relative comonadicity results proved earlier ${ }^{23}$ apply to the category of internal presheaves over $T \mathbb{C}$, rather than $\mathbb{C}$. So the direct comparison, in the relative case, is $T_{T \mathbb{C}} \sim \tilde{S}_{\mathbb{C}}$ rather than $T_{\mathbb{C}} \sim \tilde{S}_{\mathbb{C}}$. In order to obtain the latter, we embed $\left(\mathcal{N} / C_{0}\right)^{\mathbf{T}_{\mathbb{C}}}$ into $\left(\mathcal{N} / T C_{0}\right)^{\mathbf{T}_{T C}}$, in a manner that is coherent with respect to $\mathbf{T}_{\mathbb{C}}$ and $\mathbf{T}_{T \mathbb{C}}$. This requires the following lemma:

[^101]Lemma 4.25. Given any internal category, $\mathbb{C}$ in $\mathcal{N}$, the $T$ functor induces an embedding of $\mathbb{C}$-presheaves into $T \mathbb{C}$-presheaves. In other words, given some internal presheaf $\left(F_{0}, \gamma_{0}, e\right)$ :

$$
\Sigma_{T d_{1}}\left(T d_{0}\right)^{*}\left(T \gamma_{0}\right)=T\left(\Sigma_{d_{1}} d_{0}^{*}\left(\gamma_{0}\right)\right)
$$

Proof. T preserves finite limits, and is full and faithful, so it preserves all necessary structure.

As $T$ induces an embedding of $\left(\mathcal{N} / C_{0}\right)^{\mathbf{T}_{\mathbb{C}}}$ into $\left(\mathcal{N} / T C_{0}\right)^{\mathbf{T}_{T \mathbb{C}}}$, the following result can be proven as a specific case of Corollary 2.38, for the general functor defined in Proposition 2.37 (which is, in this case, Theorem 4.24). It is also worth sketching the set theoretic proof.

Proposition 4.26. Given an internal category $\mathbb{C}$ in $\mathcal{N}$, there is an embedding of the category of algebras $\left(\mathcal{N} / C_{0}\right)^{\boldsymbol{T}_{\mathbb{C}}}$, where $\boldsymbol{T}_{\mathbb{C}}=\Sigma_{d_{1}} d_{0}{ }^{*}$, into the category of relative coalgebras $\left(\mathcal{N} / T C_{0}\right)^{\tilde{S}_{\mathrm{C}}}$, defined above.

Proof. Given an internal presheaf $\left(F_{0}, \gamma_{0}, e\right)$ and the corresponding algebra:

$$
e: \Sigma_{d_{1}} d_{0}{ }^{*}\left(\gamma_{0}\right) \rightarrow \gamma_{0}
$$

one can define a morphism $\tilde{e}$ of $\mathcal{N} / T C_{0}$ :

$$
\begin{array}{lr}
\tilde{e}: T \gamma_{0} \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}\left(\gamma_{0}\right) ;\{x\} \mapsto\langle\tilde{e}(x),\{c\}\rangle & \text { (action of } \tilde{e}) \\
\tilde{e}(x): d_{0}^{-1}(c) \rightarrow C_{1} \times_{d_{1}} F ; f \mapsto\langle f, e(f, x)\rangle & \text { (action of } \tilde{e}(x))
\end{array}
$$

While the definition of $\tilde{e}$ in the relative case is the direct (type adjusted) analogue to the classical case, the resulting map is not a coalgebra. It can, however, be used to define a relative coalgebra:

$$
\chi:\left|T \gamma_{0} \downarrow T_{\mathbb{C}}\right| \rightarrow\left|T \gamma_{0} \downarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}\right|
$$

As $T$ is full and faithful, given a map $g: T \gamma_{0} \rightarrow T \beta$ in $\mathcal{N} / T C_{0}$, there is a unique map $h=T^{-1}(g): \gamma_{0} \rightarrow \beta$ in $\mathcal{N} / C_{0}$. The relative coalgebra $\chi$ is defined by the action:

$$
\chi: g \mapsto \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(h) \circ \tilde{e}
$$

$$
(\text { where } T(h)=g)
$$

Explicitly, $\chi(g): T \gamma_{0} \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}(\beta)$ is the map defined by the action:

$$
\begin{array}{ll}
\chi(g):\{x\} \in\left(T \gamma_{0}\right)^{-1}\{c\} \mapsto\langle h(\tilde{e}(x)),\{c\}\rangle & \text { (action of } \chi(g)) \\
h(\tilde{e}(x)): f \mapsto\langle f, h(e(f, x))\rangle & \text { (action of } h \circ \tilde{e})
\end{array}
$$

Corollary 4.27. The embedding defines an equivalence of categories between $\left(\mathcal{N} / C_{0}\right)^{T_{\mathrm{C}}}$ and the full subcategory of $\left(\mathcal{N} / T C_{0}\right)^{\tilde{S}_{\mathrm{C}}}$, defined by the relative coalgebras, $(\chi, T X)$, with base objects in the image of $T$.

Proof. See Proposition 2.40.

## Interpreting the $\tilde{S}_{\mathbb{C}}$-Coalgebras

The full subcategory of $\left(\mathcal{N} / T C_{0}\right)^{\tilde{S}_{\mathbb{C}}}$, formed by the embedding of $\left(\mathcal{N} / C_{0}\right)^{\mathbf{T}_{\mathbb{C}}}$ into $\left(\mathcal{N} / T C_{0}\right)^{\mathbf{T}_{T \mathbb{C}}}$, is what we would have expected. But there are presheaves $\left(F_{0}, \gamma_{0}, e\right)$ over $T \mathbb{C}$, where $F_{0}$ is not isomorphic to any object in the image of $T$. By Proposition 2.37, these correspond to relative $\tilde{S}_{\mathbb{C}}$-coalgebras, as well.

If we are interested in formally associating the relative coalgebras of $\tilde{S}_{\mathbb{C}}$ with the category of internal presheaves over $\mathbb{C}$, we need a precise understanding of the relative coalgebras that are not associated with algebraically presented $\mathbb{C}$-presheaves (i.e. $\mathbf{T}_{\mathbb{C}}$-algebras). These fall into two cases:

1. Relative coalgebras that are determined by $\mathbf{T}_{T \mathbb{C}}$-algebras.
2. Relative coalgebras that have no canonical association to $\mathbf{T}_{T \mathbb{C}}$-algebras.

Case 1: In the first case, we consider an arbitrary internal presheaf over $T \mathbb{C}$, $\left(F_{0}, \gamma_{0}, e\right)$. If $F_{0} \cong T G$, for some object $G$, we could form the relative coalgebra $(\chi, T G)$ as in Proposition 4.26. The algebra map $e: \Sigma_{T d_{1}}\left(T d_{0}\right)^{*}\left(\gamma_{0}\right) \rightarrow \gamma_{0}$ is then $\chi\left(i d_{T G}\right)$.

If $F_{0}$ is not in the image of $T$ (up to isomorphism), one can still form an associated relative coalgebra. ${ }^{24}$ To understand why the relative coalgebra can be constructed in NF, from the perspective of set theory, it is helpful to consider what would happen if one attempted to derive a standard coalgebra map $F_{0} \rightarrow \tilde{\Pi}_{d_{0}} d_{1}{ }^{*}$ from $e: \Sigma_{T d_{1}}\left(T d_{0}\right)^{*}\left(\gamma_{0}\right) \rightarrow \gamma_{0}$. This would fail for the same reason it always does. Given some $\{c\} \in T C_{0}$, the ordered pairs of the following set abstract are inhomogeneous:

$$
\left\{\langle\{c\}, h\rangle \mid h:\left(T d_{0}\right)^{-1}(\{c\}) \rightarrow T C_{1} \times_{T d_{1}} F_{0}\right\}
$$

We can subvert this issue for the case where $F_{0}=T G$, as the algebra map would also be in the image of $T$. A more general way would be to consider the "subset" of the unstratified set abstract above, defined for maps $k:\left(T d_{0}\right)^{-1}(\{c\}) \rightarrow T C_{1} \times_{T d_{1}} F_{0}$, where the image of $k$ is the same size as a set of singletons. Effectively, the relative coalgebra operation derived from $\left(F_{0}, \gamma_{0}, e\right)$ is stratified, because it respects exactly this condition.

Given a map $g: \gamma_{0} \rightarrow T \beta$ in $\mathcal{N} / T C_{0}$, the composite $g \circ e$ is defined by the diagram:


The map $\chi(g): \gamma_{0} \rightarrow \tilde{\Pi}_{d_{0}}\left(d_{1}\right)^{*}(\beta)$ is then just defined by the action:

$$
x \in \gamma_{0}^{-1}(\{c\}) \mapsto\langle\{c\}, \cup(g \circ e)\rangle
$$

Case 2: For a (standard) comonad $S$, the conditions satisfied by the object map of a Manes coalgebra $(\chi, X), \chi:\left|X \downarrow 1_{\mathcal{E}}\right| \rightarrow|X \downarrow S|$, both reduce to and arise from a corresponding classical coalgebra, $\chi\left(i d_{X}\right): X \rightarrow S X$.

For a relative comonad $S$ along $F$, the identity morphism will only exist as an object in the comma category $X \downarrow F$, when $X$ is in the image of $J$ iteslf. However, there is no reason why $X \downarrow F$ could not have an initial object, regardless of whether or not $X$ is in the image of $F$. In the standard case, where $F=1_{\mathcal{E}}$, the identity morphism $i d_{X}$ is

[^102]precisely the initial object of $X \downarrow 1_{\mathcal{E}}{ }^{25}$ Therefore, it seems that the best approximation of an (algebraically presented) presheaf over $T \mathbb{C}$ from a $\tilde{S}_{\mathbb{C}}$-relative coalgebra is some (possibly weak) form of an initial object/colimit.

If $F_{0}$ is not the size of a set of singletons, it would be particularly interesting if, for a given $T \mathbb{C}$-presheaf, one was able to derive a "best approximation," among $\mathbb{C}$-presheaves, from the corresponding relative coalgebra.

### 4.4 An Internal Relative KZ-Monad in $\mathcal{N}$

### 4.4.1 Externalisation and Fam

Definition 4.28. Given an internal category $\mathbb{C}$ in a category $\mathcal{E}$, the externalisation of $\mathbb{C}, e \mathbb{C}$, is defined as follows:

Objects: The class of objects $|e \mathbb{C}|$ is $\left|\mathcal{E} / C_{0}\right|$, where $f: I \rightarrow C_{0}$ is interpreted as an $I$-indexed family of objects in $\mathbb{C}$.

Morphisms: The morphisms of $e \mathbb{C}$ correspond to maps between families of objects:

$$
\begin{array}{lr}
(h, \alpha):\left(C_{i}\right)_{I} \rightarrow\left(D_{k}\right)_{K} & \\
\alpha: I \rightarrow K & \text { (re-indexing morphism) } \\
h_{i}: C_{i} \rightarrow D_{\alpha(i)} & \text { (I-indexed family of morphisms) }
\end{array}
$$

Internally, such a map, $(h, \alpha): f \rightarrow g$ in $e \mathbb{C}$, corresponds to the following diagram, where top and lower right triangles commute:


[^103]The externalization of an internal category $\mathbb{C}$ corresponds to the (external) coproduct completion of $\mathbb{C}, \operatorname{Fam} \mathbb{C}$. We will speak of $e \mathbb{C}$ and $\operatorname{Fam} \mathbb{C}$ interchangeably, where it is intuitive to do so ${ }^{26}$

Remark (Motivating the General Study of Indexing and Collections). When $\mathcal{E}=$ Set, the family fibration, $\delta_{\mathbb{C}}: F a m \mathbb{C} \rightarrow$ Set, is a motivating example for fibred category theory - the general study of the role "collections" play in mathematics [3, 5]. Just as $\delta_{\mathcal{C}}$ maps an indexed family of $\mathcal{C}$-objects to its underlying index, $\left(C_{k}\right)_{K} \mapsto K$, a general fibration $P: \mathcal{A} \rightarrow \mathcal{B}$ associates objects of $\mathcal{A}$ with an "indexing object" in $\mathcal{B}$. As such, one can abstract properties involving (unstructured) ${ }^{27}$ sets (e.g. local smallness, coproduct completeness, etc.) to elementary category theory. Indexed sums and products (modulo coherence conditions) are, respectively, left and right adjoint to re-indexing.

Studying the (pseudo)functor, Fam : CAT $\rightarrow C A T$, as opposed to the codomain fibration at $F a m \mathcal{C}$, is motivated by an alternative generalization: the study of free cocompletions with respect to small, discrete diagrams (i.e. free coproduct completion).

Free cocompletions, with respect to a given class of diagrams, correspond to KZ-monads [26]. There are a number of definitions one might find but, given a pseudomonad $\langle P, y, m\rangle$, the key piece of data is: the (2-dimensional) unit law is witnessed by $\alpha$ : $1_{P} \Rightarrow m P y$ and $\varepsilon: y_{P} m \Rightarrow 1_{P}$

where the natural isomorphisms $\alpha$ and $\varepsilon$ are the unit and co-unit of an adjoint string ${ }^{28}$

$$
P y \dashv m \dashv y_{P}
$$

As the pseudomonadic unit condition: $m y_{P} \cong 1_{P}$ is witnessed by the co-unit $\varepsilon$, free algebras are uniquely classified as left adjoint to the unit 1-cell $y_{P}$. Free algebras provide

[^104]what Kock refers to as a "syntactic or term model" of cocompletion. For the KZ-monad:
$$
F a m \iota \dashv \Sigma_{\text {Fam }} \dashv \iota_{\text {Fam }}
$$
this corresponds to the "syntactic" coproduct functor $\Sigma_{\text {FamC }}$ :


The free coproduct completion, FamC, satisfies the universal property: Any functor $F: \mathcal{C} \rightarrow \mathcal{D}$ between $\mathcal{C}$ and a coproduct-complete category $\mathcal{D}$ has a unique (coproduct preserving) factorization, $F^{*}: F a m \mathcal{C} \rightarrow \mathcal{D}$, given by the left extension of $F$ along $\iota_{\mathcal{C}}$.


The left extension property corresponds to the natural bijection:

$$
[F a m \mathcal{C}, \mathcal{D}]\left(F^{*}, G\right) \cong[\mathcal{C}, \mathcal{D}]\left(F, G \circ \iota_{\mathcal{C}}\right)
$$

The adjoint relationship classifying free algebras classifies all algebras. For a coproductcomplete category $\mathcal{D}$, the explicit coproduct functor $\Sigma_{\mathcal{D}}$ arises as a left adjoint to the unit $\iota_{\mathcal{D}}: \mathcal{D} \rightarrow$ FamD [26].

$$
\mathcal{D}\left(\Sigma_{\mathcal{D}}(A), D\right) \cong \operatorname{FamD}\left(A, \iota_{\mathcal{D}}(D)\right)
$$



Notice, coproduct completeness has been classified in two ways. The first diagram expressed coproduct completeness by the property that any functor $F$ with codomain $\mathcal{D}$ has a left extension along the unit of the KZ-monad. The second diagram described (functorial) coproduct completeness as being left adjoint to the unit. The former is referred to as the no-iteration classification 64].

### 4.4.2 Relative Pseudomonads

Hyland et al. have recently introduced the 2-categorical notion of a relative monad, a relative pseudomonad [8]. As with the 1-dimensional case, the relative structure is a direct generalization of the "no-iteration" presentation of (pseudo)monads. The 2dimensional no-iteration presentation of pseudomonads and pseudoalgebras is given in [38. For the relative case, we adopt the notation of [38], but the abbreviated coherence conditions of [8, Definition 3.1] ${ }^{29}$

Definition 4.29. A relative pseudomonad $D$ along $J: \mathcal{X} \rightarrow \mathcal{A}$ consists of the following data:

1. An object function $D: \mathcal{X} \rightarrow \mathcal{A}$.
2. A collection of 1-cells $i_{X}: J X \rightarrow D X$ in $\mathcal{A}$, indexed by objects of $\mathcal{X}$, which we refer to as the unit of the relative pseudomonad.
3. A collection of functors $(-)^{D}{ }_{X, Y}: \mathcal{A}[J X, D Y] \rightarrow \mathcal{A}[D X, D Y]$.
4. A natural family of invertible 2-cells $D_{(-)}$, for each morphism $F: J X \rightarrow D Y$ in $\mathcal{A}$ :

5. A natural family of invertible 2-cells $D_{(-,-)}$, for each pair of morphisms $F: J X \rightarrow$ $D Y$ and $G: J Z \rightarrow D X$ in $\mathcal{A}:$

6. A family of 2-isomorphisms $D_{X}: i_{X}{ }^{D} \rightarrow 1_{D X}$, indexed by objects of $\mathcal{X}$.
[^105]Such that, for morphisms $F: J X \rightarrow D Y, G: J Z \rightarrow D X, H: J W \rightarrow D Z$ and the pasting diagrams below, the following axioms hold:

$$
\begin{equation*}
\text { Diagram 4.1 }=1_{i_{X}} \tag{Unit}
\end{equation*}
$$

Diagram4.2 $=1_{F^{D}}$
Diagram $4.3=$ Diagram 4.4
(Associativity)




Remark (Yoneda Structures vs. Relative KZ-Monads). One of the motivations for this definition is the presheaf construction for small categories, $\mathbf{P}: \mathcal{C} \mapsto \widehat{\mathcal{C}}$. Restricting the domain of $\mathbf{P}$ to small categories allows one to consider its role as a free cocompletion.

But, size issues prevent the formation of a KZ-monad ${ }^{30}$

In order to obtain closure, in terms of size, we must work in $C A T$. In this context, however, $\mathbf{P}$ is no longer a free cocompletion. This gives rise to Yoneda Structures and the idea of "admissibility" [65, 69.

Rather than weaken the cocompletion condition as one does with Yoneda Structures, relative pseudomonads weaken the "size" condition (i.e. they work with $\mathbf{P}: C a t \rightarrow C A T$ ). As $\mathbf{P}$ is the free cocompletion functor for small categories, we recover a relative $K Z$ (pseudo)monad ${ }^{31}$ Hyland et al. refer to this as a lax idempotent relative pseudomonad [8].

Definition 4.30. A relative pseudomonad $D$ over $J: \mathcal{X} \rightarrow \mathcal{A}$ is a lax idempotent relative pseudomonad if, given any $F: J X \rightarrow D Y, D_{F}: F \rightarrow F^{D} i_{X}$ exhibits $F^{D}: D X \rightarrow D Y$ as a left extension of $F$ along the unit $i_{X}$. In other words, given any $H: D X \rightarrow D Y$ :

$$
\mathcal{A}[J X, D Y]\left(F, H \circ i_{X}\right) \cong \mathcal{A}[D X, D Y]\left(F^{D}, H\right)
$$

In addition, $(-)^{D}$ must respect left extensions:

$$
F^{D} \circ D_{G}=\left(D_{F, G} \circ i_{Z}\right) \cdot D_{F^{D} G}
$$



The unit and associativity axioms of Definition 4.29 are implied by the conditions of Definition $4.30{ }^{32}$

[^106]
### 4.4.3 Internalizing Fam in $\mathcal{N}$

In NF, even externalizations are internal:

$$
\forall \mathbb{C} \in \operatorname{cat}(\mathcal{N}) \cdot \operatorname{Fam} \mathbb{C} \in \operatorname{cat}(\mathcal{N})
$$

We can define a functor between (internal) categories $\operatorname{Fam}: \operatorname{cat}(\mathcal{N}) \rightarrow \operatorname{cat}(\mathcal{N})$, where $\operatorname{Fam} \mathbb{C}=e \mathbb{C}$. But, Fam: $\operatorname{cat}(\mathcal{N}) \rightarrow \operatorname{cat}(\mathcal{N})$ is an external functor. The appropriate internal version of Fam requires a further restriction, to obtain homogeneity.

Definition 4.31. $\tilde{e}: \operatorname{cat}(\mathcal{T N}) \rightarrow \operatorname{cat}(\mathcal{N})$ maps each internal category $T \mathbb{C}$ to the (internal) free coproduct completion: $e \mathbb{C}$. Internal functors, $T F: T \mathbb{C} \rightarrow T \mathbb{D}$, are mapped to functors $\tilde{e} F: F a m \mathbb{C} \rightarrow F a m \mathbb{D}$, determined by their actions on $(\mathbb{C})_{1}{ }^{33}$

From $\tilde{e}$, we obtain a relative lax idempotent pseudomonad.

As $\tilde{e}$ is part of a $T$-relative algebraic structure, Fam $\mathbb{C}$ is a free $T$-indexed coproduct completion of $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$. To prove this, we first show that Fam $\mathbb{C}$ is $T$-coproduct complete.

Lemma 4.32. For any $\mathbb{C}$ in $\operatorname{cat}(\mathcal{N})$, Fam $\mathbb{C}$ has $T$-indexed (internal) coproducts.

Proof. Consider an indexed family of objects, $\left((C)_{I_{j}}\right)_{J}$ in $F a m \mathbb{C}$, corresponding to a map $J \rightarrow \mathcal{N} / C_{0}$, where $\alpha_{j}: I_{j} \rightarrow C_{0}$ corresponds to the image of $j \in J$. Furthermore, assume that there exists some $K$ such that $T K \cong J$. The coproduct of $\left((C)_{I_{j}}\right)_{J}$ is defined:

$$
\coprod\left((C)_{I_{j}}\right)_{J}=(C)_{\amalg_{T K} I_{\{k\}}}=\left\{\left\langle\langle i, k\rangle, \alpha_{\{k\}}(i)\right\rangle \mid i \in I_{\{k\}}\right\}
$$

It remains to confirm the universal property. Given some $J$-indexed family of maps $\left\langle f, u_{j}\right\rangle:(C)_{I_{j}} \rightarrow(B)_{H}$, we can construct a canonical map:

$$
(f, u)_{T K}:(C)_{\amalg_{T K} I_{\{k\}}} \rightarrow(B)_{H}
$$

defined by the following action:

$$
\left\langle\langle i, k\rangle, \alpha_{\{k\}}(i)\right\rangle \mapsto f_{i, k}: C_{i, k} \rightarrow B_{u_{\{k\}}(i)}
$$

[^107]where $f_{i, k}$ is a morphism in the $I_{\{k\}}$-indexed family:
$$
\left\langle f, u_{\{k\}}\right\rangle:(C)_{I_{\{k\}}} \rightarrow(B)_{H} ; u_{\{k\}}: I_{\{k\}} \rightarrow H
$$

Each indexed family is isomorphic to a $T$-indexed coproduct of singleton families.

$$
(C)_{I} \cong \coprod_{T I}\left(C_{i}\right)_{1}
$$

Theorem 4.33. e forms an internal lax idempotent relative pseudomonad, along the inclusion $\operatorname{cat}(T \mathcal{N}) \hookrightarrow \operatorname{cat}(\mathcal{N})$.

Proof. Given an internal category $\mathbb{A} \in \operatorname{cat}(\mathcal{N})$, we can define an obvious relative unit functor:

$$
\iota_{\mathbb{A}}: T \mathbb{A} \rightarrow F a m \mathbb{A}:\{A\} \mapsto(A)_{1}
$$

Given any functor $F: T \mathbb{A} \rightarrow F a m \mathbb{D}$, there exists a left extension $F^{*}: F a m \mathbb{A} \rightarrow F a m \mathbb{D}$, presented by a natural isomorphism $\Psi_{F}: F \rightarrow F^{*} \iota_{\mathbb{A}}$.


As a convention, we denote the indexed family of $\mathbb{D}$ objects $F(\{A\})$ by the map:

$$
F(\{A\})=(D)_{I_{A}}=\gamma_{A}: I_{A} \rightarrow D_{0}
$$

To each $\{f\}:\{A\} \rightarrow\left\{A^{\prime}\right\}$ in $T \mathbb{A}$, we associate a map $F(\{f\})$ over $\alpha_{f}$ in $F a m \mathbb{D}$, denoted:

$$
\left\langle F_{f}, \alpha_{f}\right\rangle:(D)_{I_{A}} \rightarrow(D)_{I_{A^{\prime}}}
$$

Internally, $\left\langle F_{f}, \alpha_{f}\right\rangle$ is displayed by the diagram:


We define $F^{*}$ by the following action on a $K$-indexed family of objects in $A_{0}$ :

$$
F^{*}\left((A)_{K}\right) \mapsto \coprod_{T K} F\left(\left\{A_{k}\right\}\right) \equiv \gamma_{K}: \coprod_{T K} I_{A_{k}} \rightarrow D_{0}=\left\{\left\langle\langle k, i\rangle, \gamma_{A_{k}}(i)\right\rangle \mid i \in I_{A_{k}}\right\}
$$

Internally, $\coprod_{T K} F\left(\left\{A_{k}\right\}\right)$ is displayed by the morphism:

$$
\begin{aligned}
& \coprod_{T K} I_{A_{\{k\}}} \longrightarrow D_{0} \\
& \quad\langle k, i\rangle \longmapsto \gamma_{K}(k, i) \equiv \gamma_{A_{k}}(i)
\end{aligned}
$$

where $\gamma_{K}$ is the unique factorization of $\left(\gamma_{\{k\}}\right)_{T K}$ through the coproduct. ${ }^{34}$
The action of $F^{*}$ on morphisms is induced by the the action of $F$, defined above.

$$
\langle f, \beta\rangle:(A)_{K} \rightarrow\left(A^{\prime}\right)_{K^{\prime}} \mapsto F^{*}(f, \beta): \coprod_{T K} F\left(\left\{A_{k}\right\}\right) \rightarrow \coprod_{T K^{\prime}} F\left(\left\{A_{k^{\prime}}^{\prime}\right\}\right)
$$

The action of $F^{*}(f, \beta)$ is induced by the image (in $F a m \mathbb{A}$ ) of each $f(k) \in A_{1}$ under $F: T \mathbb{A} \rightarrow F a m \mathbb{A}$. This is displayed by the diagram:

$F_{f}^{*}(k, i)=F_{f(k)}(i) \in D_{1}$ determines a morphism in $\mathbb{D}$.

$$
F_{f(k)}(i): \gamma_{A_{k}}(i) \rightarrow \gamma_{A_{\beta(k)}}\left(\alpha_{f_{k}}(i)\right)
$$

The natural isomorphism $\Psi_{F}: F \rightarrow F^{*} \iota_{\mathbb{A}}$ is just the canonical bijection $X \cong(\{*\} \times X) \underbrace{[35}$

$$
\begin{aligned}
& F(\{A\}) \equiv \gamma_{A}: I_{A} \rightarrow D_{0}=\left\{\left\langle i, \gamma_{A}(i)\right\rangle \mid i \in I_{A}\right\} \\
& \iota_{\mathbb{A}}(A)=(A)_{1} \equiv!_{A}: 1 \rightarrow A_{0} \\
& F^{*} \circ \iota_{\mathbb{A}}(A) \equiv \gamma_{A}: \coprod_{T 1} I_{A_{\{*\}}} \rightarrow D_{0}=\left\{\left\langle\langle i, *\rangle, \gamma_{A}(i)\right\rangle \mid i \in I_{A}\right\}
\end{aligned}
$$

[^108]The natural isomorphism satisfying the multiplication condition, $\mu_{G, F}:\left(G^{*} F\right)^{*} \rightarrow G^{*} F^{*}$, is essentially the condition that (stratified) coproducts satisfy an associative property:

$$
\coprod_{T K} \coprod_{T I} A_{i} \cong \coprod_{T\left(\coprod_{T K} I_{k}\right)} A_{i, k}
$$

The following diagrams of $\left(G^{*} F\right)^{*}$ and $G^{*} F^{*}$ work this out, explicitly, for $F: T \mathbb{A} \rightarrow$ FamD and $G: T \mathbb{D} \rightarrow F a m \mathbb{C}$. We denote $F(\{A\})$ and $G(\{D\})$ by $\gamma_{A}: I_{A} \rightarrow D_{0}$ and $\delta_{D}: J_{D} \rightarrow C_{0}$, respectively:


$$
(K \rightarrow A) \longmapsto \coprod_{T K}\left(\gamma_{A_{k}}: I_{A_{k}} \rightarrow D_{0}\right) \longmapsto \coprod_{T\left(\amalg_{T K} I_{A_{k}}\right)}\left(\delta_{\gamma(k, i)}: J_{\gamma(k, i)} \rightarrow C_{0}\right)
$$

where the graph of the unique factorization is:

$$
\left(\delta_{\amalg_{T K} I_{A_{k}}}: \coprod_{T\left(\amalg_{T K} I_{A_{k}}\right)} J_{\gamma(k, i)} \rightarrow C_{0}\right)=\left\{\langle\langle\langle k, i\rangle, j\rangle, \delta(\gamma(k, i), j)\rangle \mid i \in I_{A_{k}} \wedge j \in J_{\gamma(k, i)}\right\}
$$

On the other hand:


As stratified coproducts are distributive in $\mathcal{N}$, it is straightforward to show the unique factorization is isomorphic to $\delta_{\amalg_{T K} I_{A_{k}}}$ in $\mathcal{N} / C_{0}$. Their respective actions are:

$$
\begin{align*}
\langle\langle k, i\rangle, j\rangle & \mapsto \delta(\gamma(k, i), j)  \tag{*}\\
\langle k,\langle i, j\rangle\rangle & \mapsto \delta(\gamma(k, i), j) \tag{*}
\end{align*}
$$

The final component in the proof is that $\Psi_{F}$ presents $F^{*}$ as a left extension of $F$ along $\iota_{\mathbb{A}}$. In other words, we want to define the adjoint $(-)^{*} \dashv-\circ \iota_{\mathbb{A}}$ :

$$
\begin{aligned}
& {[F a m \mathbb{A}, F a m \mathbb{D}]\left(F^{*}, G\right) \cong[T \mathbb{A}, F a m \mathbb{D}]\left(F, G \circ \iota_{\mathbb{A}}\right)}
\end{aligned}
$$

In a coproduct completion, $\operatorname{Fam} \mathbb{A}$, each indexed family $(A)_{K}$ is isomorphic to the coproduct of its components. In particular, we obtain a canonical factorization:

$$
\widehat{\left(!_{A_{k}}\right)_{K}}: \coprod_{T K}\left(A_{k}\right)_{1} \rightarrow(A)_{K}
$$

where $!_{A_{k}}:\left(A_{k}\right)_{1} \rightarrow(A)_{K}$ corresponds to the injection of each member into the indexed family ${ }^{36}$

$$
!_{A_{k}}=\left(i d_{A_{k}}, k\right):\left(A_{k}\right)_{1} \rightarrow(A)_{K} ; i d_{A_{k}}: A_{k} \rightarrow A_{k} ; k: 1 \rightarrow K
$$

With this in mind, consider the unit and co-unit of $(-)^{*} \dashv-\circ \iota_{\mathbb{A}}$ :

$$
\begin{align*}
& \Psi_{F}: F \rightarrow F^{*} \circ \iota_{\mathbb{A}} ; \Psi_{F, A}: F\{A\} \rightarrow \coprod_{1}(F A)_{1}  \tag{unit}\\
& \varepsilon_{G}:\left(G \circ \iota_{\mathbb{A}}\right)^{*} \rightarrow G ; \varepsilon_{G,(A)_{K}}: \coprod_{T K} G\left(\left(A_{k}\right)_{1}\right) \rightarrow G\left((A)_{K}\right) \tag{co-unit}
\end{align*}
$$

The unit is just the presentation of the left extension $\Psi_{F}: F \rightarrow F^{*} \circ \iota_{\mathbb{A}}$. The co-unit is the unique factorization of the images of the injection maps $G\left(!_{A_{k}}\right): G\left(\left(A_{k}\right)_{1}\right) \rightarrow G\left((A)_{K}\right)$, through the coproduct $\coprod_{T K} G\left(\left(A_{k}\right)_{1}\right)$ :

$$
\varepsilon_{G,(A)_{K}}=\left(\widehat{G\left(!_{A_{k}}\right)}\right)_{K}: \coprod_{T K} G\left(\left(A_{k}\right)_{1}\right) \rightarrow G\left((A)_{K}\right)
$$

For the triangle identities, we consider a pair of natural transformations:

$$
\begin{aligned}
& \tau: F^{*} \rightarrow G \\
& \vartheta: F \rightarrow G \circ \iota_{\mathbb{A}}
\end{aligned}
$$

and look to prove:

$$
\begin{aligned}
& \varepsilon_{G} \cdot\left(\tau_{\iota_{\mathbb{A}}} \cdot \Psi_{F}\right)^{*}=\tau \\
& \left(\left(\varepsilon_{G} \cdot \vartheta^{*}\right) \circ \iota_{\mathbb{A}}\right) \cdot \Psi_{F}=\vartheta
\end{aligned}
$$

The second identity is straightforward:

$$
\varepsilon_{G,(A)_{K}} \cdot \vartheta_{(A)_{K}}^{*}=\left(\widehat{G\left(!_{A_{k}}\right)}\right)_{K} \cdot \coprod_{T K} \vartheta_{A_{k}}: \coprod_{T K} F\left\{A_{k}\right\} \rightarrow \coprod_{T K} G\left(\left(A_{k}\right)_{1}\right) \rightarrow G\left((A)_{K}\right)
$$

Restricted to singleton families by horizontal precomposition with $\iota_{\mathbb{A}}$ (and vertically precomposed with $\Psi_{F}$ ), this clearly yields the original transformation, $\vartheta$, as expressed by the following diagram:

[^109]$\left(\left(\epsilon_{G} \cdot \vartheta^{*}\right) \circ \iota_{\mathbb{A}}\right) \cdot \Psi_{F}=\vartheta:$
\[

$$
\begin{aligned}
& F\{A\} \xrightarrow{\Psi_{F, A}} \coprod_{T 1} F\{A\} \xrightarrow{\vartheta_{(A)_{1}}^{*}} \coprod_{T 1} G\left((A)_{1}\right) \xrightarrow{\varepsilon_{G,(A)}} G\left((A)_{1}\right)
\end{aligned}
$$
\]

The first triangle identity is slightly less apparent. Each component of $\tau$ :

$$
\tau_{(A)_{K}}: \coprod_{T K} F\left\{A_{k}\right\} \rightarrow G\left((A)_{K}\right)
$$

corresponds to the unique coproduct factorization of a $K$-indexed family of "sub-component" maps:

$$
\tau_{(A)_{K}}^{k}: F\left\{A_{k}\right\} \rightarrow G\left((A)_{K}\right)
$$

Furthermore, naturality implies that any component $\tau_{(A)_{K}}$ yields $K$ commutative diagrams, corresponding to each $!_{A_{k}}:\left(A_{k}\right)_{1} \rightarrow(A)_{K}$.


Thus, $\tau_{(A)_{K}}$ is the unique factorization of the $K$-indexed family of maps, $\left(G\left(!_{A_{k}}\right) \circ \tau_{\left(A_{k}\right)_{1}}\right)_{K}$, through $\coprod_{T K} F\left\{A_{k}\right\}$. But, we previously defined $\tau_{(A)_{K}}$ as the unique factorization of the $K$-indexed family of "sub-component" maps $\tau_{(A)_{K}}^{k}$. Therefore, for an arbitrary family $(A)_{K}$ :

$$
\tau_{(A)_{K}}^{k}=G\left(!_{A_{k}}\right) \circ \tau_{\left(A_{k}\right)_{1}} \circ i n_{F\left\{A_{k}\right\}}
$$

As $\varepsilon_{G,(A)_{K}}$ is $\left(\widehat{G\left(!_{A_{k}}\right)}\right)_{K}$, the unique factorization of the $K$-indexed family $\left(G\left(!_{A_{k}}\right)\right)_{K}$ through the coproduct, we obtain the following commutative diagram, witnessing the desired triangle identity.
$\epsilon_{G} \cdot\left(\tau_{\iota_{\mathrm{A}}} \cdot \Psi_{F}\right)^{*}=\tau:$


The basic 2-categorical unit and associativity rules do not vary substantively from the classical case ${ }^{37}$

### 4.5 Cocomplete Objects and Relative KZ-Algebras

Algebras for a KZ-monad are left adjoint to components of the unit of the monad [26]. The content of this is: KZ-algebras provide a canonical colimiting operation (from a class of diagrams) to the base category of the algebra. One hopes for a similar classification result for relative KZ-algebras. However, the relationship between relative pseudoalgebras and free structures is no less complicated than the 1-dimensional case (see Section 2.4). Relative EM-(pseudo)algebras lack free presentations.
[8] looks at the Kleisli categories associated with a variety of relative KZ-monads. These can be thought of as the free relative pseudoalgebras. For our purposes, however, it is necessary to study the broader relative Eilenberg-Moore category of (possibly non-free) algebras.

### 4.5.1 Relative Pseudoalgebras

Definition 4.34. A relative pseudoalgebra $A$ for a $J$-relative pseudomonad $D: \mathcal{X} \rightarrow \mathcal{A}$ consists of the following data:

1. An object $A$ in $\mathcal{A}$
2. A family of functors, indexed by objects of $\mathcal{X}$,

$$
(-)^{A}: \mathcal{A}[J X, A] \rightarrow \mathcal{A}[D X, A]
$$

3. A natural family of invertible 2-cell $A_{(-)}$, for each $X$ in $\mathcal{X}$ and each morphism

[^110]$$
F: J X \rightarrow A:
$$

4. A natural family of invertible 2 -cells $A_{(-,-)}$, for each pair of objects $X, Y$ in $\mathcal{X}$ and morphisms $F: J X \rightarrow A$ and $G: J Y \rightarrow D X:$


Such that, for morphisms $F: J X \rightarrow A, G: J Y \rightarrow D X, H: J Z \rightarrow D Y$, the unit and associativity axioms hold for the following diagrams:

$$
\begin{align*}
& \text { Diagram } 4.5=1_{F^{A}}  \tag{Unit}\\
& \text { Diagram } 4.6=\text { Diagram } 4.7
\end{align*}
$$



Note: Definition 4.34 does not say anything about a left extension. We should therefore not assume that (non-free) relative pseudoalgebras of a lax idempotent relative pseudomonad inherit this property automatically. Indeed, the standard extension of the proof that any (pseudo)algebra of a KZ-Doctrine arises as an adjoint to the unit to the no-iteration case does not extend to the relative case (see Lemma 4.44).

## Cocomplete Objects and Kan Pseudoalgebras

Definition 4.35. 69] Given a KZ-monad $D$ on a 2 -category $\mathcal{C}$, an object $A$ is $D$ cocomplete if, for every $F: X \rightarrow A$, the left extension $\chi(F)$ of $F$ along the unit $\iota_{X}$ exists and is exhibited by a 2 -isomorphism $\eta_{F}: F \rightarrow \chi(F) \circ \iota_{X}$ in such a way that it respects free extensions, in the definition of the KZ-monad.

$$
\chi(F) \circ D_{G}=\eta_{\chi(F) \circ G}
$$



It is apparent that the a $D$-cocomplete object coincides with the definition of a noiteration Kan algebra, recorded as the objects of the category $\bar{D}-$ Alg in [37, Section 3]. ${ }^{38}$ The generalization of Definition 4.35 to a relative $D$-cocomplete object is straightforward. Definition 4.36. Given a relative KZ-monad $D$ along $J: \mathcal{X} \rightarrow \mathcal{A}$, an object $A \in \mathcal{A}$ and an operation $\chi:|J \downarrow A| \rightarrow|D \downarrow A|$ is relative $D$-cocomplete if the following conditions hold:

1. For each 1-cell $F: J X \rightarrow A$, a natural isomorphism $\eta_{F}: F \rightarrow \chi(F) \circ \iota_{X}$ which exhibits $\chi(F)$ as a left extension.

[^111]2. $\chi$ respects the left extensions induced by $D$. For each pair of 1-cells $F: J X \rightarrow A$ and $G: J Z \rightarrow D X$, a natural isomorphism:
$$
\beta_{F, G}: \chi(\chi(F) \circ G) \rightarrow \chi(F) \circ G^{D}
$$

A relative $D$-cocomplete object forms the base of a canonical relative pseudoalgebra, just as it does in the non-relative case.

Lemma 4.37. A relative $D$-cocomplete object $A$ in $\mathcal{A}$ is the base of a relative pseudoalgebra $A$. The correspondence is defined below, given $F: J X \rightarrow A$ and $G: J B \rightarrow D X$ :

$$
\begin{aligned}
& F^{A} \equiv \chi(F) \\
& A_{F} \equiv \eta_{F} \\
& A_{F, G} \equiv \beta_{F, G}
\end{aligned}
$$

(Left extension of $F$ along $\iota_{A}$ )
(Presentation of the left extension)
(Associativity 2-isomorphism)

Proof. Effectively, this is a corollary of $(v)$ implies $(i)$ in [8, Theorem 5.3] and (one direction) of the relative version of [37, Theorem 5.1]. ${ }^{39}$

Corollary 4.38. For any $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, Fam $\mathbb{C}$ is a relative Fam-cocomplete object. By implication, the free $T$-coproduct completion of $\mathbb{C}$ and the left extension operation:

$$
(-)_{-, \mathbb{C}}^{*}:[T(-), F a m \mathbb{C}] \rightarrow[F a m(-), F a m \mathbb{C}]
$$

form a free relative Fam-pseudoalgebra.

Proof. Theorem 4.33 and Lemma 4.37.

We wish to extend this beyond the free case, to state that $T$-coproduct complete categories correspond to Fam-cocomplete objects and, hence, relative Fam-pseudoalgebras.

Proposition 4.39. Fam is a free coproduct completion, in the sense that any $T$ coproduct complete category $\mathbb{D}$ is a relative Fam-algebra, where $\chi_{\mathbb{D}}:|T \downarrow \mathbb{D}| \rightarrow \mid F a m \downarrow$ $\mathbb{D} \mid$ is a left extension.

[^112]Proof. The proof largely mirrors Theorem 4.33 but, rather than using the properties of indexed families, we need to define a left extension along $\iota$, using only the abstract (categorical) properties given by an arbitrary small $T$-coproduct complete category $\mathbb{D}$. We can define $\chi_{\mathbb{D}}:|T \downarrow \mathbb{D}| \rightarrow|F a m \downarrow \mathbb{D}|$ as the left extension along $\iota$ in the same manner as we defined $(-)^{*}$. Given a functor $F: T \mathbb{A} \rightarrow \mathbb{D}$, we define $\chi_{\mathbb{D}}(F): F a m \mathbb{A} \rightarrow \mathbb{D}$ by the following action:

$$
\chi_{\mathbb{D}}(F):(A)_{K} \mapsto \sum_{T K} F\left(\left\{A_{k}\right\}\right)
$$

Given $\langle f, \alpha\rangle:(A)_{K} \rightarrow\left(A^{\prime}\right)_{K^{\prime}}$, we obtain the map

$$
\chi_{\mathbb{D}}(f, \alpha): \sum_{T K} F\left(\left\{A_{k}\right\}\right) \rightarrow \sum_{T K^{\prime}} F\left(\left\{A_{k^{\prime}}\right\}\right)
$$

induced by the $K$-indexed family of morphisms, where $i n_{A_{k}}$ denotes the injection of $A_{k}$ into the $\sum_{T K} A_{k}$ :

$$
F\left(\left\{A_{k}\right\}\right) \xrightarrow{F T f_{k}} F\left(\left\{A_{\alpha(k)}^{\prime}\right\}\right) \xrightarrow{\left.i n_{F\left\{A_{\alpha(k)}^{\prime}\right.}\right\}} \sum_{T K^{\prime}} F\left(\left\{A_{k^{\prime}}^{\prime}\right\}\right)
$$

Functoriality is inherited from $F$. Importantly, this includes the property $F\left(i d_{A_{k}}\right)=$ $i d_{F A}$, which implies:

$$
F^{*}\left(!_{A_{k}}\right) \circ i n_{f\left\{A_{k}\right\}}:\left(\left\langle i d_{A_{k}}, k\right\rangle:\left(A_{k}\right)_{1} \rightarrow\left(A_{k}\right)_{K}\right) \mapsto i n_{F\left\{A_{k}\right\}}
$$

where $i n_{F\left\{A_{k}\right\}}$ is the inclusion of $F\left(\left\{A_{k}\right\}\right)$ into the coproduct, $\sum_{T K} F\left(\left\{A_{k}\right\}\right.$. Thus, given a functor $G: F a m \mathbb{A} \rightarrow \mathbb{D}$ and a natural transformation $\tau: \chi(F) \rightarrow G$, we obtain a more general version of the naturality diagram, used to prove the left extension property, in Theorem 4.33.


Also following Theorem 4.33, $\eta_{F}: F \rightarrow \chi_{\mathbb{D}}(F) \circ \iota_{\mathbb{A}}$ corresponds to the canonical isomorphism $\sum_{T 1} F\left(\left\{A_{1}\right\}\right) \cong F\left(A_{1}\right)$. Also proven as in Theorem 4.33, $\eta_{F}$ presents the left extension of $F$ along $\iota_{\mathbb{A}}$.

Finally, there exists a natural isomorphism $\beta_{F, G}: \chi(\chi(F) \circ G) \rightarrow \chi(F) \circ G^{*}$.

$\beta_{F, G}$ is the canonical 2-cell between the isomorphism $\eta_{\chi(F) \circ G}: \chi(F) \circ G \rightarrow \chi(\chi(F) \circ G) \circ \iota_{\mathbb{Z}}$, exhibiting the left extension of $\chi(F) \circ G$ along $\iota_{\mathbb{Z}}$, and the isomorphism $\chi(F) \circ \Psi_{G}$ : $\chi(F) \circ G \rightarrow \chi(F) \circ G^{*} \circ \iota_{\mathbb{Z}}$. The 2-isomorphism is unique, by the universal property of left extensions.

While $F a m \mathbb{C}$ is (only) a free $T$-coproduct completion, it is a broadly appropriate notion for the classification of general coproduct complete categories in $\mathcal{N}$. In the standard case (i.e. non-relative), a coproduct complete category $\mathbb{C}$ is equivalent to a Fam -algebra. The (canonical) coproduct functor is given by the left extension of $1_{\mathbb{C}}$ along $\iota_{\mathbb{C}}$, which we refer to as $\sum_{\mathbb{C}}$.

$$
\begin{gathered}
\mathbb{C}\left(\sum_{\mathbb{C}}(C)_{K}, C^{\prime}\right) \cong \operatorname{Fam} \mathbb{C}\left((C)_{K},\left(C^{\prime}\right)_{1}\right) \\
\mathbb{C} \xrightarrow[\imath_{\mathbb{C}}]{\imath_{\mathbb{C}}} \operatorname{Fam} \mathbb{C} \\
\underbrace{\mid \Sigma_{\mathbb{C}}}_{\mathbb{C}}
\end{gathered}
$$

In the relative case, such a canonical (internal) coproduct functor is not available in general. If we take categories of the form $T \mathbb{C}$, however, the left extension property of a relative Fam-cocomplete object furnishes a "literal" coproduct functor, which we refer to as $\widetilde{\Sigma}_{\mathbb{C}}$ :


As above, we obtain a natural isomorphism:

$$
T \mathbb{C}\left(\widetilde{\sum_{\mathbb{C}}}(C)_{K},\left\{C^{\prime}\right\}\right) \cong \operatorname{Fam} \mathbb{C}\left((C)_{K},\left(C^{\prime}\right)_{1}\right)
$$

Any $K$-indexed family of maps in $\mathbb{C}$ to a singleton family $\left(C^{\prime}\right)_{1}$, corresponds to unique map from $\widetilde{\sum}_{\mathbb{C}}\left((C)_{K}\right)$ in $T \mathbb{C}$ to $\left\{C^{\prime}\right\}$. As $T$ is a full and faithful embedding, we obtain
a unique correspondence (although not a stratified, internal isomorphism):

$$
T \mathbb{C}\left(\widetilde{\sum_{\mathbb{C}}}(C)_{K},\left\{C^{\prime}\right\}\right) \sim \mathbb{C}\left(\sum_{\mathbb{C}}(C)_{K}, C^{\prime}\right)
$$

where $\sum_{\mathbb{C}}(C)_{K}=T^{-1}\left(\widetilde{\sum}_{\mathbb{C}}(C)_{K}\right)$ clearly satisfies the universal property of a coproduct in $\mathbb{C}$. Similarly, the relevant family of $T \mathbb{C}$-maps, corresponding to $\left(f_{k}: C_{k} \rightarrow C^{\prime}\right)_{K}$ is the $T K$-indexed family:

$$
\left(\{f\}_{k}:\left\{C_{k}\right\} \rightarrow\left\{C^{\prime}\right\}\right)_{T K}
$$

The $T K$-indexed family of maps $(\{f\})_{T K}$ factors uniquely through $\widetilde{\sum}_{\mathbb{C}}(C)_{K}$, by the adjoint transpose $(\tilde{f})_{K}$ of $(f)_{K}$. Thus, we obtain the result:

Lemma 4.40. An internal category $T \mathbb{C} \in \operatorname{cat}(\mathcal{N})$ is a relative $\tilde{e}$-cocomplete object (and so a pseudoalgebra) if and only if $T \mathbb{C}$ is a $T$-coproduct complete category. Furthermore, we obtain a stratified, internal coproduct functor as the left extension of $1_{T \mathbb{C}}$ along $\iota_{\mathbb{C}}$.

Lemma 4.41. For any $T \mathbb{C}$ in $\operatorname{cat}(T \mathcal{N}), T \mathbb{C}$ has $T$-indexed coproducts if and only if $\mathbb{C}$ has all (internal) coproducts.

Proof. $T$ is a full and faithful embedding.

Together, these results prove:
Corollary 4.42. $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$ is coproduct complete if and only if $T \mathbb{C}$ is a relative e-cocomplete object.

While we don't obtain an internal coproduct functor, we can associate a canonical coproduct with any family of objects and a unique factorization of any family of maps:

$$
(f)_{K}:\left(C_{k} \rightarrow C^{\prime}\right)_{K} \sim \cup(\tilde{f})_{K}: \bigcup \widetilde{\sum_{\mathbb{C}}}(C)_{K} \rightarrow C^{\prime}
$$

The internal lax idempotent relative pseudomonad $\tilde{e}$, while only $T$-coproduct complete, classifies the coproduct complete internal categories of $\mathcal{N}$.

However, as is the case with relative $\tilde{S}_{\mathbb{C}}$-coalgebras, there is no obvious reason a category $\mathbb{D}$, not in the image of $T$, could not be a relative $\tilde{e}$-cocomplete object. In this instance,
the converse of Proposition 4.39 is less straightforward: is a relative e é-cocomplete object, not in the image of $T$, a coproduct complete category?

Furthermore, it is not clear that the standard equivalence between KZ-algebras and cocomplete objects extends to the relative case. In particular, the coherence conditions required to prove the converse of Lemma 4.37 (in the standard case) rely on the existence of a canonical 2-isomorphism, which has no clear analogue in the relative case.

### 4.5.2 Are All Relative KZ-algebras Cocomplete?

Here we investigate the general question of whether or not the equivalence between $P$ cocomplete objects (i.e. Kan algebras) and standard pseudoalgebras (for a KZ-monad) extends to lax idempotent relative pseudomonads.

Lemma 4.37 gives one direction: The proof that Kan (pseudo)algebras (what we have referred to as $P$-cocomplete objects) are pseudoalgebras extends to the relative case. What remains open is whether the classification result for "standard" pesudoalgebras of a KZ-monad carries over. In other words: We seek to determine whether the left extension property in the definition of a relative KZ-monad (and its Kleisli algebras) implies that any generic relative EM-algebra corresponds to a left extension along the relative unit.

## Normalized Case

For "normalized" algebras, the unit condition holds up to strict identity, rather than isomorphism. For a (pseudo)monad $(T, y, m)$, where unit and multiplication form the following adjoint string:

$$
T y \dashv m \dashv y_{T}
$$

[26] classifies a KZ-monad by a series of conditions that must be satisfied, where $\lambda$ denotes a modification $\lambda: T y \Rightarrow y_{T}{ }^{40}$

$$
\begin{align*}
& m \circ T y=m \circ y_{T}=i d_{T}  \tag{T0}\\
& \lambda \circ y=1_{T y \circ y}=1_{y_{T} \circ y}  \tag{T1}\\
& m \circ \lambda=1_{m \circ T y}=1_{m \circ y_{T}}  \tag{T2}\\
& m \circ T m \circ \lambda_{T}=1_{m \circ T m \circ T y_{T}}=1_{m \circ T m \circ y_{T^{2}}}=1_{m} \tag{T3}
\end{align*}
$$

From these conditions, it follows that $m \dashv y_{T}$ has an identity co-unit. The unit is given by $T m \circ \lambda_{T}$ [26, Proposition 1.2]. The dual result holds for $T y \dashv m$ [26, Proposition 3.1].

Definition 4.43. [26, Definition 2.1] A map, $a: T A \rightarrow A$, which is a reflection left adjoint to the unit (i.e. $a \dashv y_{A}$, with an identity co-unit) is referred to as a structure.

The classification theorem for KZ-monads states an equivalence between classical $T$ (pseudo)algebras and structures. Given a normalized $T$-(pseudo)algebra ( $A, a, \chi$ ), the first condition (the identity co-unit) of what Kock refers to as a "structure" holds automatically $\left(a \circ y_{A}=1_{A}\right)$. The unit of $a \dashv y_{A}$ is then given by:

$$
T a \circ \lambda_{A}: T a \circ T y_{A}=i d_{T A} \rightarrow T a \circ y_{T A}=y_{A} \circ a
$$

Proving $(A, a, \chi)$ is equivalent to a structure reduces to proving the triangle identities for the unit and (identity) co-unit. The first identity holds as a direct corollary to condition $T 1$.

$$
T a \circ \lambda_{A} \circ y_{A}=1_{y_{A}}
$$

The second identity:

$$
a \circ T a \circ \lambda_{A}=1_{a}
$$

relies on the coherence condition for multiplication, satisfied by the pseudoalgebra $(A, a, \chi) \cdot \sqrt[41]{\square}$

$$
\chi \circ y_{T A}=\chi \circ T y_{A}=1_{a}
$$

[^113]One then proves the triangle identity by composing the following diagram in the two ways, given by the exchange law for vertical and horizontal composition.


Remark (No-Iteration KZ-Monads). The "normalized" proof extends to (normalized) no-iteration pseudoalgebras, as a special case of the more general results in [37], where $T$ "need not be an endofunctor." One should note, however, the use of $\lambda=\eta \circ T y$ does not extend to the relative case. Even avoiding the adjoint string by stating, simply, that $T_{(-)}$presents a natural family of left extensions, we are forced to confront the fact that, given any $A, 1_{T A}$ need not be an object of the comma category $J \downarrow T$ and, therefore, $T_{1_{T A}}$ may not exist in the first place, to be used as it is in [37, Theorem 4.1].

## More General Result: The Lax Case

The more general equivalence, for the lax case, also relies on coherence conditions for classical pseudoalgebras [36, Section 10]. As we are relaxing the strictness condition on the (isomorphic) unit and co-unit, $\alpha: 1_{T} \rightarrow m T y$ and $\varepsilon: m y_{T} \rightarrow 1_{T}$, the appropriate definition of a KZ-Doctrine requires the coherence condition:

$$
(\varepsilon \cdot \alpha) \circ y=m \circ\left(y_{y}\right)^{-1}
$$

where $y_{y}$ is the 2-cell induced by the natural transformation $y .42$ This results in a lax version of condition T1:

$$
\begin{equation*}
\lambda \circ y=y_{y} \tag{*}
\end{equation*}
$$

The identity $\lambda=\eta \circ T y$ is now:

$$
\lambda=\left(y_{T} \circ \alpha^{-1}\right) \cdot(\eta \circ T y)
$$

In [26], Kock proves an equivalence between normalized pseudoalgebras and "structures." Similarly, in the lax case, one seeks to prove a pseudoalgebra ( $X, x, \sigma, \chi$ ) corresponds to

[^114]an adjoint $x \dashv y_{X}$. The unit 2-isomorphism $\sigma: x \circ y_{X} \rightarrow 1_{X}$ turns out to be the co-unit of $x \dashv y_{X}{ }^{43}$ From $\sigma$ one can construct a 2 -cell $\hat{\sigma}$ as in Diagram 4.8. $\hat{\sigma}$ is the unit of $x \dashv y_{X}$.


As in the normalized case, the first of the two triangle identities is corollary to condition $T 1$ (in this case, $T 1^{*}$ ). The second identity corresponds to proving the pasting Diagram 4.9 is equivalent to the 2-identity $1_{x}$. Once again, this requires using pseudoalgebraic coherence conditions.


Taking the approach of [37] one can prove that " $\sigma$ is the co-unit of $x \dashv y_{X}$ " is equivalent to the result: " $\sigma^{-1}$ presents a 'generating' left extension." In other words, ( $X, x, \sigma, \chi$ ) generates a no-iteration pseudoalgebra where, given any $H: Z \rightarrow X, X_{H}: H \Rightarrow H^{X} \circ y_{Z}$ presents $H^{X}$ as a left extension, defined as the pasting:


[^115]Given a 2-cell $\tau: H \rightarrow K \circ y_{Z}$, one can prove the universal property of a left (Kan) extension (i.e. the family of adjunctions $\left.(-)_{Z}^{X} \dashv-\circ y_{Z}\right)$ by defining the adjoint transpose of $\tau, \hat{\tau}: H^{X} \rightarrow K$, similar to Diagram 4.9.


The proof that $\hat{\tau}$ pasted along the right edge of Diagram 4.10 is equivalent to $\tau$ depends on condition $T 1^{*} \stackrel{44}{4}$ Proving uniqueness of $\hat{\tau}$ requires the coherence conditions of the pseudoalgebra.

## "Pure" No-Iteration

As outlined above, [37] proves an equivalence between no-iteration pseudomonads with the (Kan) extension property and lax idempotent pseudomonads. But our goal is to determine whether or not this proof extends to the relative case. Relative pseudomonads generalize no-iteration pseudomonads, but do not have corresponding classical pseudomonads.

Therefore, to see if the classification theorem of KZ-monads extends to the relative case, we need to carry out a "pure" no-iteration proof: Assuming only the coherence conditions of a no-iteration pseudomonad (see [38]) and that $\left(T_{(-)},(-)^{T}, y\right)$ possesses the left extension property, we wish to prove each no-iteration pseudoalgebra $\left(X_{(-)},(-)^{X}\right)$ presents a left extension.

Diagram 4.10 determines $\hat{\tau}$, starting from a lax idempotent pseudomonad, so we wish to construct a "pure" no-iteration version. The first step is to determine the no-iteration

[^116]version of the modification $\lambda$.

Using the equivalence:

$$
\begin{equation*}
\lambda=\left(\beta \circ y_{T}\right) \cdot\left(T y \circ \varepsilon^{-1}\right) \tag{4.12}
\end{equation*}
$$

we can replace $\beta$ and $\varepsilon$ according to the correspondences of [37, Theorem 4.1].

$$
\begin{aligned}
& \varepsilon_{Z} \equiv T_{1_{T Z}}^{-1} \\
& \beta_{Z} \equiv \xlongequal{1_{y_{T Z} \circ y_{Z}}} \cdot T_{y_{T Z} \circ y_{Z}, 1_{T Z}}^{-1}
\end{aligned}
$$

Both instances use the correspondence $m=1_{T-}^{T}$ and the latter equivalence uses two iterations of the left extension property of the no-iteration pseudomonad:

$$
\begin{align*}
& 1_{y_{T Z} \circ y_{Z}}: y_{T Z} \circ y_{Z} \rightarrow y_{T Z} \circ y_{Z} \\
& \sim \widehat{1_{y_{T Z} \circ y_{Z}}}:\left(y_{T Z} \circ y_{Z}\right)^{T} \rightarrow y_{T Z}  \tag{1}\\
& =\widehat{1_{y_{T Z} \circ y_{Z}}}:\left(y_{T Z} \circ y_{Z}\right)^{T} \rightarrow 1_{T^{2} Z} \circ y_{T Z} \\
& \sim \widehat{1_{y_{T Z} \circ y_{Z}}}:\left(\left(y_{T Z} \circ y_{Z}\right)^{T}\right)^{T} \rightarrow 1_{T^{2} Z} \tag{2}
\end{align*}
$$

Thus we can re-write Formula 4.12 as the pasting:


Using [38, Definition 2.1, Axiom 5 ${ }^{45}$ and that $T_{y_{T Z} \circ y_{Z}}$ presents a left extension we obtain the result:

$$
\begin{equation*}
\lambda_{Z}=\widehat{1_{y_{T Z} \circ y_{Z}}} \tag{4.14}
\end{equation*}
$$

The correspondence between classical and no-iteration pseudomonads, given in [38], and Formula 4.14 allow one to give a "pure" no-iteration version of Diagram 4.10.

[^117]

In the relative case, a pseudoalgebra $\left(X_{(-)},(-)^{X}\right)$ does not have a canonical "generating" extension $\left(X_{1_{X}}\right)$ as in the classical case. Using the coherence conditions of no-iteration pseudomonads, however, we can re-write Diagram4.15 as Diagram4.16, without assuming the existence of a generating extension $X_{1_{X}}: 1_{X} \rightarrow\left(1_{X}\right)^{X} \circ y_{X}$.


On the other hand, viewing $\hat{\tau}$ through the lens of the adjunction defining the left extension property,

$$
\vartheta^{Z, X}, \varphi^{Z, X}:(-)_{Z}^{X} \dashv-\circ y_{Z}
$$

we can give an equivalent definition of $\hat{\tau}$ using the co-unit $\varphi^{Z, X}$ :

$$
\hat{\tau}=\varphi_{K}^{Z, X} \cdot \tau^{X}
$$

Thus, we claim that Diagram4.16 is the vertical composite of $\tau^{X}$ with the co-unit. This claim is proven in the following lemma. Lemma 4.44 proves the no-iteration pseudoal-
gebra, $\left(X_{(-)},(-)^{X}\right)$, inherits the left extension property from $\left(T_{(-)},(-)^{T}, y\right)$, by proving the following transformations satisfy adjoint triangle conditions:

$$
\begin{align*}
& \vartheta^{Z, X}=X_{(-)}  \tag{unit}\\
& \varphi_{K}^{Z, X}=X_{K}^{-1} \cdot\left(K^{X} \circ \widehat{1_{y_{T Z} \circ y_{Z}}}\right) \cdot X_{K, y_{T Z} \circ y_{Z}} \cdot\left(X_{K} \circ y_{Z}\right)^{X} \tag{co-unit}
\end{align*}
$$

Lemma 4.44. Given any no-iteration pseudoalgebra $X$ and any 1-cell $K: T Z \rightarrow X$, we can define the component of the left extension co-unit at $K, \varphi_{K}^{Z, X}$, as:

$$
\begin{equation*}
\varphi_{K}^{Z, X}=X_{K}^{-1} \cdot\left(K^{X} \circ \widehat{1_{y_{T Z} \circ y_{Z}}}\right) \cdot X_{K, y_{T Z} \circ y_{Z}} \cdot\left(X_{K} \circ y_{Z}\right)^{X} \tag{4.17}
\end{equation*}
$$



Proof. As a co-unit, $\varphi_{K}^{Z, X}$ must be equivalent to $\widehat{1_{K \circ y_{Z}}}$. Thus, using only the properties of a no-iteration pseudoalgebra, we must prove that 4.17 pasted onto the diagram displaying $X_{K \circ y_{Z}}$ is the unique 2-cell, yielding $1_{\text {Koy }_{Z}}$.

The relevant pasting, written as a vertical composition of 2-cells, is the first triangle identity of the left extension adjunction $(-)^{X} \dashv-\circ y_{Z}$ :

$$
\begin{equation*}
\left(\left(X_{K}^{-1} \cdot\left(K^{X} \circ \widehat{1_{y_{T Z} \circ y_{Z}}}\right) \cdot X_{K, y_{T Z} \circ y_{Z}} \cdot\left(X_{K} \circ y_{Z}\right)^{X}\right) \circ y_{Z}\right) \cdot X_{K \circ y_{Z}} \tag{4.18}
\end{equation*}
$$

Using naturality of $X_{(-)}$, we obtain:

$$
\left(\left(X_{K} \circ y_{Z}\right)^{X} \circ y_{Z}\right) \cdot X_{K \circ y_{Z}}=X_{K^{X} \circ y_{T Z} \circ y_{Z}} \cdot\left(X_{K} \circ y_{Z}\right)
$$

Then, using [38, Definition 4.1, Axiom 3]:

$$
\left(X_{K, y_{T Z} \circ y_{Z}} \circ y_{Z}\right) \cdot X_{K^{x} \circ y_{T Z} \circ y_{Z}}=K^{X} \circ T_{y_{T Z} \circ y_{Z}}
$$

Therefore we can rewrite Formula 4.18 as:

$$
\left(\left(X_{K}^{-1} \cdot\left(K^{X} \circ \widehat{1_{y_{T Z} \circ y_{Z}}}\right)\right) \circ y_{Z}\right) \cdot\left(K^{X} \circ T_{y_{T Z} \circ y_{Z}}\right) \cdot\left(X_{K} \circ y_{Z}\right)
$$

As $T_{y_{T Z} \circ y_{Z}}$ presents a left extension:

$$
\left(\widehat{1_{y_{T Z} \circ y_{Z}}} \circ y_{Z}\right) \cdot T_{y_{T Z} \circ y_{Z}}=1_{y_{T Z} \circ y_{Z}}
$$

Thus, Formula 4.18 reduces to:

$$
\left(X_{K}^{-1} \cdot X_{K}\right) \circ y_{Z}=1_{K \circ y_{Z}}
$$

For the second triangle identity, we need to prove:

$$
\varphi_{H^{X}}^{Z, X} \cdot\left(X_{H}\right)^{X}=1_{H^{X}}
$$



Using [38, Definition 4.1, Axiom 3], replace the instances of $X_{H^{X}}$ and $X_{H^{X}}^{-1}$ with the following identities:

$$
\begin{aligned}
& X_{H^{X}}=\left(X_{1_{T Z}, H} \circ y_{T Z}\right)^{-1} \cdot\left(H^{X} \circ T_{1_{T Z}}\right) \\
& X_{H^{X}}^{-1}=\left(H^{X} \circ T_{1_{T Z}}\right)^{-1} \cdot\left(X_{1_{T Z}, H} \circ y_{T Z}\right)
\end{aligned}
$$

Thus, Diagram 4.19 becomes:


We can replace the upper right half of the main rectangle, by using [38, Definition 4.1, Axiom 6]:


Using this identity, we take the vertical composite:

$$
\left(\left(X_{1_{T Z}, H} \cdot X_{1_{T Z}, H}^{-1}\right) \circ y_{T Z} \circ y_{Z}\right)^{X} \cdot\left(H^{X} \cdot T_{1_{T Z}} \circ y_{Z}\right)^{X}=\left(H^{X} \cdot T_{1_{T Z}} \circ y_{Z}\right)^{X}
$$

Using the left extension property of $T_{(-)}$, we can also take the following identity, where
 by $T_{1_{T Z}^{T} \circ y_{T Z} \circ y_{Z}}$.

$$
\left(1_{T Z}^{T} \circ \widehat{1_{y_{T Z} \circ y_{Z}}}\right) \cdot T_{y_{T Z} \circ y_{Z}, 1_{T Z}}=\widehat{1_{1_{T Z}^{T} \circ y_{T Z} \circ y_{Z}}}
$$

We can now re-write Diagram 4.19 again:


The following identity is given by [38, Definition 4.1, Axiom 4]:


By making this replacement and using an instance of [38, Definition 4.1, Axiom 2], proving that Diagram 4.19 $=1_{H^{x}}$ is equivalent to proving Diagram 4.20 $=T_{Z}$ (the unique transpose of $1_{y_{z}}$, induced by the "free" left extension $\left.T_{(-)}\right)$.


As $T_{(-)}$has the left extension property, $T_{Z}$ is the unique 2 -cell corresponding to $1_{y_{Z}}$. Thus, proving that the pasting of $T_{y z}$ to the top of Diagram 4.20 results in the identity $1_{y_{z}}$ is equivalent to proving our desired result: Diagram 4.20 $=T_{Z}$. But the former follows almost immediately. As $(-)^{T}$ is functorial, we have the identity:

$$
\left(\left(T_{1_{T Z}} \circ y_{Z}\right)^{T} \circ y_{Z}\right) \cdot T_{y_{Z}}=T_{1_{T Z}^{T} \circ y_{T Z} \circ y_{Z}} \cdot\left(T_{1_{T Z}} \circ y_{Z}\right)
$$

From here the middle cells cancel (by definition of $\widehat{1_{1_{T Z} \circ y_{T Z} \circ y_{Z}} \text { ) and we are left with the }{ }^{T} \text {. }}$, identity:

$$
\left(T_{1_{T Z}}^{-1} \cdot T_{1_{T Z}}\right)=1_{y_{Z}}
$$

This allows us to, at last, reduce Diagram 4.19 to the diagram:


By [38, Definition 4.1, Axiom 2], the diagram is equivalent to $1_{H^{x}}$. This proves the second triangle identity and completes the lemma.

Given we have a no-iteration pseudomonad, with the left extension property for $T_{(-)}$, we state the "free" left extension co-unit as:

$$
\bar{\varphi}^{Z, T X}:\left(-\circ y_{Z}\right)^{T} \Rightarrow 1_{[T Z, T X]}
$$

As $(-)^{T}$ is functorial (between internal hom-categories), we immediately obtain

$$
\begin{equation*}
\widehat{1_{y_{T Z} \circ y_{Z}}}=\bar{\varphi}_{y_{T Z}}^{Z, T^{2} Z} \tag{4.21}
\end{equation*}
$$

Combining Formulas 4.17 and 4.21 we obtain a restatement of the previous lemma, giving an explicit description of the universal left extension property of each (no-iteration) pseudoalgebra as a composite with the co-unit $\bar{\varphi}$ witnessing the left extension property of the (no-iteration) pseudomonad.

Corollary 4.45. Given a no-iteration KZ-Doctrine, $\left(T,(-)^{T}, y\right)$, for any (no-iteration) pseudoalgebra $X$ and object $Z$, the co-unit $\varphi^{Z, X}$ of the adjunction defining the left extension property of $X_{(-)}$is determined by the co-unit $\bar{\varphi}^{Z, T^{2} Z}$ defining the "free" left extension property of $T_{(-)}$.

$$
\begin{equation*}
\varphi_{K}^{Z, X}=X_{K}^{-1} \cdot\left(K^{X} \circ \bar{\varphi}_{y_{T Z}}^{Z, T^{2} Z}\right) \cdot X_{K, y_{T Z} \circ y_{Z}} \cdot\left(X_{K} \circ y_{Z}\right)^{X} \tag{4.22}
\end{equation*}
$$

## The Relative Case

The challenge is immediate. In the case of a relative pseudomonad, we need not have any of the components in Formula 4.22. What the previous lemma has highlighted is the difficulty in moving away from a construction dependent on 1 or 2-cells that need not exist in the relative case. However, we have not yet developed anything resembling a counterexample or a necessary/sufficient condition for the relative functor ${ }^{46}$ What we ultimately desire is a construction like Diagram 4.15, which exists in a sufficiently generic (no-iteration) form that it is independent of restrictions (i.e. further conditions) in the definition of a relative lax idempotent pseudomonad in [8].

## $4.6 \tilde{\mathcal{N}}$

In the final section, we return to the study of $\operatorname{Fam}(\tilde{\mathcal{N}})$, but focus on its role as an internalization of the codomain fibration, $\mathfrak{c o d}$, rather than as a coproduct completion. We first prove that $\tilde{\mathcal{N}}$ is a full internal subcategory, generated by an object in $\mathfrak{c o d}^{-1}(V)$. This is somewhat unexpected as results from Chapter 3 show the codomain fibration cannot be locally small. ${ }^{47}$ However, despite the fact that $\mathfrak{c o d}$ is not locally small in general, the object $\pi_{2}: \in_{\mathcal{N}} \rightarrow V$ over $V$, corresponding to the stratified membership relation, is among those that do give rise to an exponential (in $\mathcal{N} / V \times V$ ).

While our original interest in $\tilde{\mathcal{N}}$ was as the naive internalization of $\mathcal{N}$ within itself, the proof that $\tilde{\mathcal{N}}$ is a full internal subcategory places it on more rigorous, categorical footing, as a candidate for an internal universe object. ${ }^{48}$ As such, the remainder of this section examines whether it is most appropriately considered as a universe of types or a universe of sets ${ }^{49}$

[^118]
### 4.6.1 $\operatorname{Fam}(\tilde{\mathcal{N}})$

Definition 4.46. A fibration $P: \mathcal{C} \rightarrow \mathcal{E}$ is locally small if, for any pair of objects $A, B$ in $\mathcal{C}$, there is an object $\Phi$ in $\mathcal{E}$, a map $\langle i, j\rangle: \Phi \rightarrow P A \times P B$, and a map $\gamma$ in the fibre $P^{-1}(\Phi)$ satisfying the universal property: Given a pair of maps $\langle f, g\rangle: X \rightarrow P A \times P B$ and a map $\alpha: f^{*} A \rightarrow g^{*} B$, there is unique map $h: X \rightarrow \Phi$ such that $h^{*} \gamma=\alpha$.


Intuitively, the map $\langle i, j\rangle: \Phi \rightarrow P A \times P B$ can be read as $\langle d o m$, cod $\rangle$, taking any map in $\mathcal{C}(A, B)$ to its source and target. The generic map $\gamma$ can be seen as a $\mathcal{C}(A, B)$-indexed family of morphisms. ${ }^{50}$ In the case of $F a m \mathcal{C}$, the generic map is the family of all possible morphisms between all possible pairs of objects in the indexed families $A$ and $B$.

$$
\coprod_{a \in P^{-1}(A)} \coprod_{b \in P^{-1}(B)} \mathcal{E}(a, b)
$$

Given a locally small fibration $P: \mathcal{C} \rightarrow \mathcal{E}$, any object $X$ in $\mathcal{C}$ induces an internal category in $\mathcal{E}$, where $P X$ is the object of objects. An internal category arising in this manner is referred to as a full internal subcategory. The full detail of this construction is covered in [23, Section B.2.3]. We define the special case of a full internal subcategory in a topos, with respect to the codomain fibration.

Definition 4.47. In a topos $\mathcal{E}$, any map $f: X \rightarrow C$ corresponds to a full internal subcategory $\mathcal{E}[f]$. The object of objects $\mathcal{E}[f]_{0}$ is $C$, the codomain of $f$. The object of morphisms is the exponential object, $\mathcal{E}[f]_{1}=\pi_{2}^{*} f_{1}^{\pi_{1}^{*} f}$, induced by local cartesian closure.

[^119]In other words, $\mathcal{E}[f]$ corresponds to the full subcategory of $\mathcal{E}$ spanned by the fibres of $f$ or, more formally, the subcategory spanned by the pullback of $f$ along each of the global elements of $C$.

Theorem 4.48. $\tilde{\mathcal{N}}$ is a full internal subcategory, generated by $\Gamma \in \mathfrak{c o d}^{-1}(V)$ :

$$
\Gamma \equiv \pi_{2}: \epsilon_{\mathcal{N}} \subset T V \times V \rightarrow V
$$

The associated exponential in $\mathcal{N} / V \times V$ is precisely the set of $N F$ functions, with domain and codomain maps:

$$
\left\langle d_{0}, d_{1}\right\rangle: \operatorname{Mor}(\mathcal{N}) \rightarrow V \times V
$$

The generic morphism:

$$
e v: d_{0}^{*} \Gamma \rightarrow d_{1}^{*} \Gamma
$$

is defined by the action: $\langle f,\{x\}\rangle \mapsto\langle f,\{f(x)\}\rangle$.

Proof. The failure of local cartesian closure in NF implies $\mathfrak{c o d}: \mathcal{N}^{2} \rightarrow \mathcal{N}$ is not locally small. Nevertheless, $\mathfrak{c o d}$ may induce a full internal subcategory on certain objects of a given fibre. In particular, we focus on $\mathfrak{c o d}^{-1}(V)$, the fibre over $V$ and the stratified set membership relation:

$$
\in_{\mathcal{N}}=\{\langle\{x\}, y\rangle \mid x \in y\}
$$

This induces an object, denoted as $\Gamma$ :

$$
\Gamma=\pi_{2}: \in_{\mathcal{N}} \rightarrow V \in \mathfrak{c o d}^{-1}(V)
$$

We can then form the pullbacks, $d_{0}^{*} \Gamma$ and $d_{1}^{*} \Gamma$, of $\Gamma$ along the respective maps $d_{0}, d_{1}$ : $\operatorname{Mor}(\mathcal{N}) \rightarrow V$. The result is a pair of morphisms in $\mathcal{N} / \operatorname{Mor}(\mathcal{N})$, which display families of objects indexed by the (homogeneous) functional relations of NF.

$$
\begin{aligned}
& d_{0}^{*} \Gamma=\{\langle f, y, z,\{x\}\rangle \mid f: y \rightarrow z \wedge x \in y\} \\
& d_{1}^{*} \Gamma=\left\{\left\langle f, y, z,\left\{x^{\prime}\right\}\right\rangle \mid f: y \rightarrow z \wedge x^{\prime} \in z\right\}
\end{aligned}
$$

From this we can define a form of evaluation map:

$$
e v: d_{0}^{*} \Gamma \rightarrow d_{1}^{*} \Gamma ;\langle f, y, z,\{x\}\rangle \mapsto\langle f, y, z,\{f(x)\}\rangle
$$

This map between members of the fibre $\mathfrak{c o d}^{-1}(\operatorname{Mor}(\mathcal{N}))$ is the proposed generic morphism. It must satisfy the following universal property: Given a pair of morphisms $f, g: W \rightarrow V$ and a map $\alpha: f^{*} \Gamma \rightarrow g^{*} \Gamma$, there is a unique map $\bar{\alpha}: W \rightarrow \operatorname{Mor}(\mathcal{N})$ such that $\bar{\alpha}^{*}(e v)=\alpha$.
$\alpha$ is defined by the following action, the nature of which allows one to associate $\alpha(\langle w,\{x\}, f(w)\rangle)$ with $\left\langle w,\left\{\hat{\alpha}_{w}(x)\right\}, g(w)\right\rangle$ :

$$
\alpha: f^{*} \Gamma \rightarrow g^{*} \Gamma ;\langle w,\{x\}, f(w)\rangle \mapsto\left\langle w,\left\{\hat{\alpha}_{w}(x)\right\}, g(w)\right\rangle
$$

Therefore, to any map $\alpha: f^{*} \Gamma \rightarrow g^{*} \Gamma$, we may associate a $W$-indexed family of maps $\hat{\alpha}_{w}: f(w) \rightarrow g(w)$. The map $\bar{\alpha}: W \rightarrow \operatorname{Mor}(\mathcal{N})$ is then defined:

$$
\bar{\alpha}: W \rightarrow \operatorname{Mor}(\mathcal{N}) ; w \mapsto \hat{\alpha}_{w}
$$

Reindexing of the proposed generic morphism ev along $\bar{\alpha}$ yields a map:

$$
\bar{\alpha}^{*}(e v): \bar{\alpha}^{*} d_{0}^{*} \Gamma=f^{*} \Gamma \rightarrow \bar{\alpha}^{*} d_{1}^{*} \Gamma=g^{*} \Gamma
$$

The action of $\bar{\alpha}^{*}(e v)$ is:

$$
\bar{\alpha}^{*}(e v):\left\langle w,\left\langle\hat{\alpha}_{w}, f(w), g(w),\{x\}\right\rangle\right\rangle \mapsto\left\langle w,\left\langle\hat{\alpha}_{w}, f(w), g(w),\left\{\hat{\alpha}_{w}(x)\right\}\right\rangle\right\rangle
$$

As $w$ uniquely determines $\left\langle\hat{\alpha}_{w}, f(w), g(w)\right\rangle$, there is a canonical choice of pullback for $d_{0}$ and $d_{1}$, respectively:

$$
\bar{\alpha}^{*} d_{i}^{*}=\{\langle w,\{x\}, f(w)\rangle \mid x \in f(w)\}
$$

Given this choice of pullback, the action of $\bar{\alpha}^{*}(e v)$ can now be restated:

$$
\langle w,\{x\}, f(w)\rangle \mapsto\left\langle w,\left\{\hat{\alpha}_{w}(x)\right\}, g(w)\right\rangle
$$

Such an action is exactly the definition, given above, for $\alpha$ itself. Furthermore, as functions in NF are extensional (i.e. two functions with the same graph are, by definition, the same function), $\bar{\alpha}$ is unique.

We investigate the full internal subcategory generated by $\epsilon_{\mathcal{N}}$ in two contexts: the role of $\tilde{\mathcal{N}}$ as a universe of sets and as a universe of types.

### 4.6.2 $\tilde{\mathcal{N}}$ As A Universe

The most common (foundational) strategy for dealing with the inherent size issues in category theory is to work within a Grothendieck Universe, $\mathcal{U}$. Essentially, $\mathcal{U}$ is a set of "small" objects satisfying certain closure and reflection properties, with respect to an ambient universe of sets.

Definition 4.49. A Grothendieck Universe is a set $\mathcal{U}$, satisfying the following properties:

1. $\mathcal{U}$ is transitive.
2. $\mathcal{U}$ is closed under the powerset operation.
3. If $x \in \mathcal{U}$ and $y \in \mathcal{U}$, then $\{x, y\} \in \mathcal{U}$.
4. If $I \in \mathcal{U}$ and $\left(X_{i}\right)_{I}$ denotes an $I$-indexed family of sets in $\mathcal{U}$, then $\bigcup_{i \in I}\left(X_{i}\right)_{I} \in \mathcal{U}$.

The "modern" conception of a Grothendieck universe is a full internal topos 51 Just as a full internal topos reflects the structure of an ambient topos, the full internal subcategory $\tilde{\mathcal{N}}$ reflects properties of $\mathcal{N}$. Extending the approach of Taylor to NF allows us to consider $\tilde{\mathcal{N}}$ and, more specifically, its generating object $\Gamma: \in_{\mathcal{N}} \rightarrow V$ from the perspective of both set theory (a universal set) and domain theory (a universal type).

## Dependent Type Theory

In [67], Taylor extends the association, objects $\sim$ types, for a category $\mathcal{C}$, to objects $\sim$ types-in-context, for the corresponding arrow category $\mathcal{C}^{2}$. Each object $X$ in $\mathcal{C}$ is still a "type" as before, but can also serve as a context for the type $\Phi: Y \rightarrow X .{ }^{52}$ Type dependency is analogous to variable declaration:

$$
\Phi: Y \rightarrow X \text { corresponds to }[\Phi, x: X] \rightarrow[x: X]
$$

[^120]Each fibre of $\Phi$ corresponds to the type obtained by substituting a term of type $X$ for the variable $x: X . Y$ is then viewed as the coproduct of an indexed family of types, dependent on the substitution of some term of type $X$. Thus:

$$
\Phi: Y \rightarrow X \text { can be viewed as } \coprod_{x \in X} \Phi[x] \rightarrow[x: X]
$$

The "elements" (i.e. terms) of type $X$ are the generalized elements of category theory, hence the type $\Phi[a]$ corresponds to the pullback diagram:


Similarly, dependent sums and products correspond to constructive existential and universal quantifiers. Thus, the extent to which dependent products exist in slice categories over given contexts in $\mathcal{C}$ is the extent to which one is able to (universally) quantify over dependent types.

The relationship between local cartesian closure and (constructive) universal quantification implies that a locally cartesian closed category (such as a topos) is the appropriate "practical" foundational setting for set/class theory. However, the field of domain theory (and the consistent admission of a universal type) implies the appropriate case is more general.

Definition 4.50. 67] The class of display maps, $\mathcal{D} \subseteq \operatorname{Arr}(\mathcal{C})$, is formed of morphisms that permit the formation of product types (i.e. formation of dependent products). A $\operatorname{map} f: Y \rightarrow X$ in $\mathcal{D}$ is denoted ' $f: Y \rightsquigarrow X$.'

Intuitively, $\mathcal{D}$ defines the subclass of $\operatorname{Arr}(\mathcal{C})$ that represents dependent types (i.e. supports variable declaration). As any model of $Z F$ forms a topos, $\mathcal{S}$, the class of display maps for dependent type theory corresponding to "classical" set theory is equivalent to $\operatorname{Arr}(\mathcal{S}), \mathcal{D}=\operatorname{Arr}(\mathcal{S})$. Just as Cantor's Theorem prevents a universal set, it precludes the existence of a universal type ${ }^{53}$

[^121]The apparent trade-off between domain theory and set theory turns out to be very familiar to those who study NF: One is forced to choose between (the possibility of) a universal, classifying type and the general formation of product types.

Ultimately, viewed as a universe, $\tilde{\mathcal{N}}$ appears to sit somewhere between a universal type and a universal set ${ }^{54}$ A version of Russell's paradox applies in both contexts, manifesting set-theoretically as the failure of (local) cartesian closure and type-theoretically as the classification ("naming") of a proper subset of the dependent types of $\mathcal{N}$.

The failure of local cartesian closure in NF, with its consistent admission of a universal set, makes $\mathcal{N}$ a tempting object of study as a model for some cousin of polymorphic lambda calculus [47]. In the context of stratified comprehension, one obtains a particularly intuitive argument for the failure of constructive universal quantification. We include the argument as an informal lemma.

Lemma 4.51 (Informal). Given a model of dependent type theory, $\mathcal{N}$, arising from a model of NF set theory, the class of display maps $\mathcal{D}$ is a proper subclass of $\operatorname{Arr}(\mathcal{N})$.

Proof. Constructive universal quantification over a given context $X$ is the assignment of a proof for each element of type $X$. Given some $\Phi$ defined in the context $X$, the type of proofs of $\forall x . \Phi[x]$ (i.e $\prod_{x \in X} \Phi[x]$ ) is a dependent type over $X$, with the fibre over each element $a: X$, corresponding to the "proofs" of $\Phi[a]$. But proofs of $\forall x . \Phi[x]$ correspond to sections of $\Phi \rightsquigarrow X$. In other words, a "proof" would be a set:

$$
\{\langle a, p(a)\rangle \mid p(a) \text { proves } \Phi[a]\}
$$

A constructive proof is a witness. Therefore, a proof in dependent type theory is a global element. Thus the syntactic definition of the set given above is:

$$
\{\langle a, p(a)\rangle \mid a: X \wedge p(a): 1 \rightarrow \Phi[a]\}
$$

Clearly, therefore, a proof of $\forall x . \Phi[x]$ is an inhomogeneous relation.

[^122]In this sense, both intuitively and formally (by way of their representation in a category), we obtain the idea 5

Proofs are a higher type than what they prove.

## Prop vs. Type

At the level of objects (i.e. working with a vanilla category $\mathcal{C}$ ), the existence of a universe object can be thought of as an object $U$, into which every object $C$ has a monomorphism. But the existence of $U$, in isolation, is not nearly as interesting without the existence of $\Omega$, a sub-object classifier. In this case, to any object $C$, we can associate a classifying map $\chi_{C}: U \rightarrow \Omega$. In effect, what is being said is: the ability of a universal object (a maximal object in the partial ordering induced by the subobjects of $\mathcal{C}$ ) to "name" the objects of $\mathcal{C}$ is dependent upon the existence of a generic classifying map, $t: 1 \rightarrow \Omega$.

Even without a universal object, however, $t: 1 \rightarrow \Omega$ is a form of local classifier, naming the subobjects of any object in $\mathcal{C}$. Viewing $\mathcal{C}$ at the level of morphisms, $\mathcal{C}^{2}$, allows us to consider $t$ as a generic object in $\mathcal{C} / \Omega$, classifying (locally) a proper subclass of the dependent types over each object $C$. In this way, we obtain a propositional classifier, Prop [67]. One can then make the important distinction, for example, between universal quantification of types (i.e. product types) and the special case of universal quantification over propositions ${ }^{56}$

Definition 4.52. A subclass $\mathcal{M} \subseteq \operatorname{Arr}(\mathcal{C})$, given a model of dependent type theory $(\mathcal{C}, \mathcal{D})$, denotes the subclass of propositional dependent types.

Example 4.53. In a topos $\mathcal{S}$, corresponding to a model of classical set theory, $\mathcal{M}$ is the class of arrows $\operatorname{Mono}(\mathcal{S})$. Accordingly, $[X \mid \Phi]$, a proposition $\Phi$ in the context $X$, corresponds to the subclass of $x: X$, for which $\Phi[x]$ is true. As the subobject classifer

[^123]$\Omega$ "classifies" the propositional types, it corresponds to Prop, the second order type of all propositional types:


Remark (Propositions and Comprehension). The identification of propositional types with monomorphisms (generalized subsets) is consistent with the axiom scheme of separation - the mapping of propositions (syntax) to types (sets). Looking at the "type" aspect of a propositional type provides a potentially deeper insight. Given a proposition $\Phi$ in the context $X$ :

$$
[X \mid \Phi]: \Phi \multimap X
$$

$[X \mid \Phi]$ displays an $X$-indexed family of objects, consisting of only $\{*\}$ and $\emptyset$. As these are the terminal and initial objects of $\mathcal{S}$, respectively, they can be seen to correspond to $\top$ (truth) and $\perp$ (falsity). Exchanging the codomain fibration $\mathfrak{c o d}$ for the more general concept of a fibred category, one can actually recover the idea of separation/comprehension as a functor $\mathfrak{C}$, right adjoint to the functor $\mathfrak{T}$, mapping each type $X$ to the terminal object $[X \mid]$ of the Lindenbaum Algebra of propositions over $X .{ }^{57}$

$$
\text { Type } \mid \text { Prop } \underset{\mathfrak{C}}{\stackrel{\mathfrak{T}}{\leftrightarrows}} \text { Type }
$$

$\mathcal{N}$ has a subobject classifier which, as in the case of $\mathcal{S}$, corresponds to Prop, with $t: 1 \rightarrow 2$ serving as the "generic" proposition in both cases. Another interesting potential for $\mathcal{N}$, however, is the existence of a type classifier.

## A Universal Type

The existence of a universal type, Type, is more than just the existence of a universe object. While $V$ is a universe object, in the sense that it classifies every object of
${ }^{57}$ Taylor goes further, to point out the generative role of separation/comprehension in Zermelo Type Theory [67]. From a fibrational standpoint, we see the set builder types $(1, \emptyset, \times, P)$, effectively, correspond to vertical maps, and set formation under separation corresponds to a sort of free closure. An investigation could be made into $K F$ Type Theory, and the role of stratified $\Delta_{0}$-comprehension as a form of closure.
$\mathcal{N}$, a type classifier requires the existence of a generic object in $\mathcal{N}^{2}$. Just as we are not interested in $\Omega$ in isolation, but rather the classifying map $t: 1 \rightarrow \Omega$; we are not interested in $V$ (i.e. Type), so much as:

$$
\Gamma: \in_{\mathcal{N}} \rightarrow V \in \mathfrak{c o d}^{-1}(V)
$$

the generating object of $\tilde{\mathcal{N}}$.

The (global) elements of Type are, in a sense, syntactic names for the semantic objects of $\mathcal{N}$. If product types (i.e. dependent products) exist in general, the classifying dependent type is a display map and we get a type theoretic analogue of Cantor's Theorem, in the same way one would for the universe of sets, $V$. The type theoretic version of Cantor's Theorem states, in effect: not all types are named by Type.

In $\mathcal{N}$, however, $\mathcal{D} \subsetneq \mathcal{N}^{2}$. Thus, it would seem that $\Gamma$ could classify ("name") all dependent types, just as the failure of cartesian closure allows for the existence of a universal set, $V$, classifying all objects. However, despite the restriction on dependent products, $\Gamma: \in_{\mathcal{N}} \rightarrow \tilde{\mathcal{N}}_{0}$ retains the appearance of what one might think of as a "set theoretic" universe - it names only a proper subclass of the dependent types of NF.

Lemma 4.54. The type classifier of $\mathcal{N}, \Gamma: \in_{\mathcal{N}} \rightarrow \tilde{\mathcal{N}}_{0}$, classifies the proper subclass of dependent types:

$$
[[-]]^{-1}(\operatorname{Fam}(\mathcal{N})) \subsetneq \operatorname{Arr}(\mathcal{N})
$$

Proof. The generalized elements of $V$, are just morphisms whose codomain is $V$. A morphism $\Phi: Y \rightarrow X$ is said to be classified by $\Gamma$ if it arises as a pullback of $\Gamma$ along some $\operatorname{map} \tau_{\Phi}: X \rightarrow V$.

$\Phi$ is (up to isomorphism) the first projection $\pi_{1}$ over $X$, of the following set:

$$
Y=X \times_{V} \in_{\mathcal{N}}=\left\{\left\langle x,\left\langle\{y\}, \tau_{\Phi}(x)\right\rangle\right\rangle \mid y \in \tau_{\Phi}(x)\right\}
$$

The fibre of $\Phi$ over any element $x \in X$ is, therefore, canonically isomorphic to $\iota$ " $\tau_{\Phi}(x)$.

Notice, this does not say that $Y$ itself is the size of a set of singletons, but rather that there is a partition of $Y$, each of component of which is the size of a set of singletons.

But the apparent weakness of $\Gamma$, expressed by Lemma 4.54, must be understood in the broader context of $\mathcal{N}$, and the ability of NF to "display" indexed families (or, in this case, dependent types) as families of fibres of maps in $\mathcal{N}$. Recall the motivating example of for general category theory in NF:

$$
\mathcal{N} / C \cong[T C, \mathcal{N}]
$$

In this context, the "set theoretic" aspect of $\Gamma$ (the naming function, $[[-]]: \mathcal{N} / V \rightarrow$ $\mathcal{N}^{2}$, is not surjective) furnishes a form of coherence for its role as a type classifier. While $\mathcal{N}^{2}$ does not internalize all indexed families ${ }^{[58}$ the named dependent types form a subcategory of $\mathcal{N}^{2}$ "displaying" a class of indexed families exactly equivalent to $\operatorname{Fam}(\mathcal{N})$.

Theorem 4.55. The universal dependent type $\Gamma: \in_{\mathcal{N}} \rightarrow \tilde{\mathcal{N}}_{0}$ classifies the objects of a subcategory of $\mathcal{N}^{2}$, (externally) equivalent to $\operatorname{Fam}(\mathcal{N})$.

Proof. As we have said previously, a map $f: Y \rightarrow X$ in $\mathcal{N}$ displays a $T X$-indexed family of sets. Therefore, any type $f: Y \rightarrow X$ in $\mathcal{N}$, which is classified by $\Gamma$, displays a $T X$-indexed family of sets of singletons. In other words, $\Gamma$ classifies those dependent types which display $\operatorname{Fam}_{T}(T \mathcal{N}){ }^{59}$ Unlike the subcategory of $\operatorname{Fam}(\mathcal{N})$ whose class of objects is those displayed by all dependent types in $\mathcal{N}$, there is an obvious equivalence of categories between $\operatorname{Fam}_{T}(T \mathcal{N})$ and $\operatorname{Fam}(\mathcal{N})$.

One might object to our emphasis on the external equivalence of categories, $\operatorname{Fam}_{T}(T \mathcal{N}) \cong$ $\operatorname{Fam}(\mathcal{N})$, when we have taken such pains to work internally, throughout Chapter 4. However, we should think back to the relative coalgebraic presentation of internal presheaves. While the comma category, $T \mathcal{N} \downarrow \mathcal{N}$, is a proper subcategory of $\operatorname{Fam}(\mathcal{N})$, it can also be seen as a cocompletion of $\operatorname{Fam}(\mathcal{N})$. Thus, despite the chain of proper

[^124]subcategories:
$$
\operatorname{Fam}_{T}(T \mathcal{N}) \subsetneq \operatorname{Fam}_{T}(\mathcal{N}) \subsetneq \operatorname{Fam}(\mathcal{N})
$$
externally, $\operatorname{Fam}_{T}(T \mathcal{N}) \cong \operatorname{Fam}(\mathcal{N})$. While one can (externally) inject $\operatorname{Fam} \mathrm{F}_{T}(\mathcal{N})$ and $\operatorname{Fam}(\mathcal{N})$ into each other, in both directions, $\operatorname{Fam}_{T}(T \mathcal{N})$ is the subcategory of $\operatorname{Fam}_{T}(\mathcal{N})$ canonically (externally) isomorphic to $\operatorname{Fam}(\mathcal{N})$.

## From Universal Type to Universal Presheaf

The connection between "named" dependent types and the subcategory of internal $\mathcal{N}$ presheaves, (externally) isomorphic to the external presheaf category, can be made more formally. Just as the category of dependent types classified by $\Gamma$ is isomorphic to $\operatorname{Fam}(\mathcal{N})$, the freely generated internal presheaf in $\mathcal{N}^{\tilde{\mathcal{N}}}$, corresponding to $\Gamma \in \mathcal{N} / V$, classifies the subcategory of internal presheaves for each internal category $\mathbb{C}, \mathcal{N}^{\mathbb{C}}$, canonically isomorphic to the category of internal functors $[\mathbb{C}, \tilde{\mathcal{N}}]$.

To this point, $T \mathcal{N}$ has denoted the full subcategory of $\mathcal{N}$ in the image of $T$. This is not precisely the same as the category resulting from applying $T$ to the internal diagram describing $\tilde{\mathcal{N}}$, the resulting such category is denoted $\mathcal{N}^{\iota}$.

Definition 4.56. We define the category $\mathcal{N}^{\iota}$ as the category presented by the $T$-image of $\tilde{\mathcal{N}}$ (i.e whose object of objects is $T V$ and whose object of maps is $T(\operatorname{Mor}(\mathcal{N})$ ), with structure maps inherited directly from $\tilde{\mathcal{N}}$ ).

The following lemma is trivial to prove in NF, but is worth noting: $\mathcal{N}^{\iota}$ is not simply isomorphic to $T \mathcal{N}, \Gamma$ displays the canonical isomorphism between them. ${ }^{60}$

Lemma 4.57. The freely generated internal presheaf of $\Gamma$ displays the isomorphism of categories $\mathcal{N}{ }^{\varrho} \cong T \mathcal{N}$.

[^125]Proof. Straightforward. Let $R_{\tilde{\mathcal{N}}}$ be the free presheaf functor defined earlier, over the internal category $\tilde{\mathcal{N}}$. The category of elements associated to $R_{\tilde{\mathcal{N}}}(\Gamma)$ is canonically isomorphic (taking the first projection) to $T \mathcal{N}$.

Theorem 4.58 (The Classifying Presheaf). Consider the cat( $\mathcal{N}$ )-indexed category $\mathcal{N}^{(-)}$, whose objects are the internal presheaf categories of each internal category in $\mathcal{N}$, and whose reindexing morphisms are induced by pullbacks along internal functors. The internal presheaf $R_{\tilde{\mathcal{N}}}(\Gamma)$ in $\mathcal{N}^{\tilde{\mathcal{N}}}$ is generic, in the sense that it classifies, for each $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, the full subcategory of internal presheaves which is canonically equivalent to the internal functor category $[\mathbb{C}, \tilde{\mathcal{N}}]$.

Proof. In one direction, consider an internal functor $F: \mathbb{C} \rightarrow \tilde{\mathcal{N}}$, given by the maps $F_{0}: C_{0} \rightarrow V$ and $F_{1}: C_{1} \rightarrow F u n$, such that the following diagram commutes:

We define $\widehat{F}_{0}$ as the pullback $C_{0} \times_{F_{0}} \in_{\mathcal{N}}$


The set corresponding to $\widehat{F}_{0}$ is defined:

$$
\widehat{F}_{0} \equiv\left\{\left\langle c,\left\langle\{x\}, F_{0}(c)\right\rangle\right\rangle \mid x \in F_{0}(c)\right\}
$$

The action map $e: C_{1} \times_{d_{0}} \widehat{F}_{0} \rightarrow \widehat{F}_{0}$ is then defined by the following action, for a map $f: c \rightarrow c^{\prime}$ in $\mathbb{C}:$

$$
\left\langle f, c,\left\langle\{x\}, F_{0}(c)\right\rangle\right\rangle \mapsto\left\langle c^{\prime},\left\langle\left\{F_{1}(f)(x)\right\}, F_{0}\left(c^{\prime}\right)\right\rangle\right\rangle
$$

In the other direction, consider an internal presheaf $\left\langle G_{0}, \gamma_{0}, e\right\rangle$ in $\mathcal{N}^{\mathbb{C}}$, whose fibres $\gamma_{0}^{-1}(c)$ are the size of sets of singletons. The pair of maps $\bar{G}_{0}$ and $\bar{G}_{1}$, corresponding internal functor $\bar{G}: \mathbb{C} \rightarrow \tilde{\mathcal{N}}$, are clearly induced by $\gamma_{0}$ and $e$.

$$
\bar{G}_{0}: C_{0} \rightarrow V ; c \mapsto \cup \gamma_{0}^{-1}(c)
$$

$$
\bar{G}_{1}: C_{1} \rightarrow F u n ;\left(f: c \rightarrow c^{\prime}\right) \mapsto \lambda x . \cup e\langle f, x\rangle: F_{0}(c) \rightarrow F_{0}\left(c^{\prime}\right)
$$

It is straightforward to show these operations are mutually inverse.

Once again, we do not obtain a classification of all internal presheaves, but rather the full subcategory displaying the internalized functor category between $\mathcal{C}$ and $\tilde{\mathcal{N}}$. In this way, the internal presheaves that are not classified by $R_{\tilde{\mathcal{N}}}$ can be seen as cocompletions.

## A (Grothendieck) Universe of Sets

As mentioned above, the appropriate generalization of a Grothendieck universe is not simply an internal topos. A Grothendieck universe of "small sets," within an ambient model of ZF set theory, is both a subcategory $\mathcal{G} \subset \mathcal{S}$ and an internal topos $\mathbb{G}$ within the topos $\mathcal{S}$ of $\mathrm{ZF}(\mathrm{C})$ sets. Thus, the appropriate generalization of is a full internal topos $\mathbb{G}$ 67].

The embedding of $\mathbb{G}$ into $\mathcal{S}$ is given by the global sections functor:

$$
U: \mathbb{G} \rightarrow \mathcal{G} \subset \mathcal{S} ; x \mapsto \mathcal{S}(1, x)
$$

The extent to which we can "access" the members of $\mathbb{G}$ is restricted to the hom-set of global elements, $\mathcal{S}\left(1, \mathbb{G}_{0}\right)$. In $\mathrm{ZF}(\mathrm{C})$, of course, there is a canonical isomorphism in $\mathcal{S}$ :

$$
\mathcal{S}\left(1, \mathbb{G}_{0}\right) \cong \mathbb{G}_{0}
$$

In NF, this is not the case ${ }^{61}$ The "small" sets, named by $\tilde{\mathcal{N}}$, form the proper embedding:

$$
\mathcal{N}(1,-): \tilde{\mathcal{N}} \mapsto \mathcal{N}
$$

Thus, while $\tilde{\mathcal{N}}$ is explicitly an internal universe of sets, it retains certain properties that are more type-theoretic in nature. It almost seems appropriate to think of a global element of $V$ as a syntactic name for the set $\iota$ " $X$.

If we conceive of the "named" sets of $\mathcal{N}$ as the "small" sets, we might consider that, rather than the strongly cantorian sets, it is those sets in the image of $T$ that are "small."

[^126]Just as the members of a Grothendieck universe can be seen as naming the small sets of an ambient model of set theory, the objects of $\tilde{\mathcal{N}}$ should not be thought of as semantic objects in their own right, but as syntactic names for small sets.

Taking $T \mathcal{N}$ as the category of "small" sets of NF, we appear to achieve some basic closure properties. The subcategory $T \mathcal{N}$ of named sets of $\mathcal{N}$ is closed under powersets in fact, the naming operation $[[-]]$ forms the canonical isomorphism $P(T X) \cong T(P X)$. One also achieves a form of closure under sumsets.

Proposition 4.59. $\tilde{\mathcal{N}}$ is self co-complete, in the sense that any family $X: I \rightarrow V$, where $I \cong U[J]$ for some $x \in V$, has a coproduct $\coprod_{I} X_{i}$ in $\tilde{\mathcal{N}}$.

Proof. Straightforward. As $I \cong U[J]$ is equivalent to $I \cong T J$, we simply form the set:

$$
\coprod_{I} X_{i}=\left\{\langle j, x\rangle \mid \exists i \in I .\{j\}=i \wedge x \in X_{i}\right\}
$$

While it may seem that we should ask for $\tilde{\mathcal{N}}$ to be closed under coproducts taken over all sets of $\mathcal{N}$, restricting to coproducts taken over the "small" sets of $\mathcal{N}$ is appropriate, in this context. ${ }^{62}$ In a similar manner, $\tilde{\mathcal{N}}$ is closed under powersets $(P(T X) \cong T(P X))$. However, these closure properties are not all they seem. In particular, there is no internal endofunctor $\tilde{P}: \tilde{\mathcal{N}} \rightarrow \tilde{\mathcal{N}}$, such that $[[-]] \circ \tilde{P}=P \circ[[-]]$.


The restriction of $P$ to the subcategory SC of strongly cantorian sets, on the other hand, is locally homogeneous (i.e. the restriction of $P$ to internal sub-categories $\mathbb{C} \subset \tilde{\mathcal{N}}$ of strongly cantorian categories can be internalized). The same, of course, is true of exponential adjunction. The trade-off being that a form of the Burali-Forti paradox has already been shown to prove $\mathbf{S C}$ cannot be internalized.

[^127]Remark (Closure vs. Size: Large Cardinals). Viewed through the lens of closure conditions, rather than cardinality, the seemingly contradictory result that the sub-class of the universal set, containing all the strongly cantorian sets, is both "smaller" and "external," becomes more clear. The phenomenon can also be seen as a further piece of "empirical" evidence, for a position advocated in [67]: large cardinals are perhaps best understood by the closure conditions of their associated internal categories.

In this way, a subcategory of $\mathcal{N}$ which is "large" (in the sense that it is external) is not determined by "size," but by its closure conditions, consistent with Taylor's comments on large cardinals. At the very least, the relationship between "external" and "closure," in the case $\mathbf{S C}$, furnishes a more semantic explanation than expressibility of a given property (such as that of being strongly cantorian) by stratified formulae.

## Chapter 5

## NF Lambda Calculus

NF has been suggested as a potential model for pure $\lambda$-calculus, at least twice in the literature. Scott observed this possibility in [58] and Forster mentioned it in 9. The temptation arises from the existence of "large" sets that, unlike most set theories, yield a straightforward interpretation of total functions as sets. But this alone does not give a combinatory algebra. The main challenge is to develop a stratified, homogeneous form of application $\sqrt{1}$ For NF, the key to constructing a combinatory algebra is not the interpretation of functions-as-sets, but that of sets-as-functions.

Scott conjectured his multi-relation model could be carried out in NF, yielding a calculus of continuous functions [55, 58. We prove Scott's conjecture ${ }^{2}$ This conjecture is stronger than it might seem. The ability of NF to model untyped lambda calculus does not make it unique among set theories ${ }^{3}$ Rather than making the assertion that a model of untyped lambda calculus can be constructed within NF, the stronger property implied by Scott's

[^128]conjecture is: any model of NF has a natural interpretation as a (multi-relation) model of untyped lambda calculus. Accordingly, the continuous functions abstracted by the model are total (i.e. are defined $V \rightarrow V$ ). The latter property is not surprising, the advantage of NF is its admission of large sets (and, importantly, a fixed point of $P$ ). But, as a trade-off, we expect to be restricted to stratified formulae. Evaluation (application, in the context of lambda calculus) should, therefore, not exist as a function in our model. However, the relevant continuous functions for a model of untyped lambda calculus are those of finite character. As a result, implementation of Scott's model requires only finite sequences, which can be coded (homogeneously) in any model of NF ${ }_{4}^{4}$ In turn, one is able to implement a homogeneous form of evaluation, sufficient to form a model of untyped lambda calculus in any model of NF.

Another consequence of homogeneous evaluation is the implementation of certain combinators, corresponding to continuous functions that are not definable (as sets) in NF. But such combinators correspond only to the curried form of some homogeneous multivariate function, whose (standard) graph is a set. Thus, we are not altering the functions of NF, just the manner in which they are presented. As such, we do not need to worry about the admission of certain inhomogeneous functions (or anti-continuous ones, such as complementation) that might allow for the construction of paradoxes.

Scott made a second conjecture, also in [58], that one could code an exact correspondence between the universe of sets (in a model of NF) and the collection of all finite sequences of sets, denoted $V^{*}$. Investigating this conjecture, we uncover certain complications arising from the implementation of sequences by iterated nesting of Quine pairs. ${ }^{5}$ We introduce an alternative construction, Quine sequences, resulting in an interpretation of NF with surjective sequences ${ }^{[6]}$ Quine sequences prove an extended version of Scott's

[^129]second conjecture:
$$
V=V^{* *} \equiv V^{\omega} \cup \bigcup_{n \in N} V^{n}
$$

The extension of multi-relations to include $\omega$-sequences leads us to consider a choice principle (obviated in the finite case) related to the $S$ combinator. This, in turn, motivates a more general study of "sequences-as-codes" in Section 5.5.

## Overview of Chapter 5

Section 5.1 gives a brief overview of Scott's original work on models of untyped $\lambda$ calculus, leading to his multi-relation model [55, 58]. The total function space $V \Rightarrow V$ forms a set in NF, and the distributive (resp. continuous) $\lambda$-operator forms a stratified mapping between $V \Rightarrow V$ and the set of $\lambda$-abstracts of distributive (resp. continuous) functions. Thus, we can consider the $\lambda$-abstracted fragment of any total function in NF.

Section 5.2 develops the algebra of distributive functions in NF. Their computational power is not particularly strong - there is a bijection between relations and distributive functions [57]. In NF, however, the calculus of distributive functions has two interesting properties. First, as $V \times V=V$, we obtain a form of $\eta$-equivalence: extensional equivalence of sets corresponds to intensional equivalence of combinators. Furthermore, we can inject the function space $V \Rightarrow V$ into the collection of distributive functions in NF, using the $j$-operator.

It is also worth looking at the category of (distributive) retracts. Scott defines the category of retracts, combinators that behave like "types" (they are idempotent), for a given combinatory algebra [57]. The subcategory consisting of sets that, as distributive combinators, behave like "types" bears a surprising resemblance to the standard category of NF sets.

In Section 5.3, we construct a multi-relation model of $\lambda$-calculus in NF. The main issue to overcome is the interpretation of sets-as-sequences. Any set in NF corresponds to a sequence of length $n$, for each $n \in N$. However, despite having the property

$$
\forall n \in N . V^{n}=V
$$

sets do not correspond to sequences of canonical length. We develop an interpretation of sets-as-sequences, where each set corresponds to a unique finite sequence and $V^{*} \subset V$. The corresponding algebra of functions confirms Scott's first conjecture, but we do not obtain an exact equivalence between $V^{*}$ and $V \square$

Section 5.4 considers the correspondence between $V$ and $V^{\omega}$, presented in [10], but forms a "quotient" $V^{\omega} / \sim_{\omega}$, corresponding to the identity $]^{8}$

$$
\left\langle x_{0}, \ldots x_{n}, \emptyset, \emptyset, \ldots\right\rangle \sim_{\omega}\left\langle x_{0}, \ldots, x_{n}\right\rangle
$$

Under this identity, sets correspond to sequences of canonical, but varying (possibly infinite) length. At this point, we are in a position to consider Scott's second conjecture. However, closer study of streams generated by recursive nesting of Quine pairs uncovers a redundancy that is not present in the coding of finite sequences. This leads us to introduce Quine sequences. As with streams derived from Quine pairs, Quine sequences can be defined recursively. But, unlike iterative nesting, Quine sequences yield: $V^{\omega}=V$. Thus, applying the identity, $\sim_{\omega}$, we obtain an exact equivalence:

$$
V=V^{* *} \equiv V^{\omega} \cup \bigcup_{n \in N} V^{n}
$$

Extending coded sequences from $V^{*}$ to $V^{* *}$ has a nontrivial impact on the resulting combinatory algebra. Certain combinators, such as $S$, are not independent of the definition of application. The appropriate notion of continuity, therefore, appears to be $\omega$-continuity, moving from functions determined by their action on finite sets to those determined by their action on countable sets. $\omega$-continuity of $S$, in particular, implies the need for a form of countable choice to construct a combinator that is $\beta$-equivalent to $\lambda x y z . x(z)(y(z))$. The dependency of continuity on the ordinal length of coded sequences leads us to consider a more general construction in Section 5.5 .

Multi-relation models of $\lambda$-calculus can be regarded as special cases of a more abstract

[^130]form of coding, expressed by the diagram:


In a multi-relation model, code : $V \rightarrow V$ is implemented by the composite action:

$$
x \mapsto \vec{x} \mapsto\left\{y \mid \exists x_{i} \in \vec{x} . y=x_{i}\right\}
$$

Any given coding operation $\chi: V \rightarrow V$ induces a canonical form of application and abstraction. $\chi$-continuous functions are those determined by their action on coded sets (i.e. those in the image of $\chi$ ) 9 Furthermore, for any given any function, one can determine its $\chi$-continuous fragment.

The ultimate goal of this section is to determine the relationship between properties of a given coding operation $\chi$ and the resulting class of $\chi$-continuous functions. Specifically, whether there are conditions on $\chi$, which are equivalent to $\beta$-equivalence between the standard combinators $S, K$ and $I$, and their $\chi$-continuous fragments. Framing the question in this way allows us to study the relationship between ( $\chi$-)continuity of $S$ and choice more formally. It also provides a better framework for understanding the combinatory algebra corresponding to the multi-relation model of NF, extended to include $\omega$-sequences ${ }^{10}$

### 5.1 Scott's Model

The first explicit description of a model of pure $\lambda$-calculus was given in 55. Specifically, for any injective $T_{0}$-space ${ }^{11}$ Scott gave an (inverse) limit construction of a topological

[^131]space, $D_{\infty}$, homeomorphic to its own function space, $\left[D_{\infty} \rightarrow D_{\infty}\right]$. A more convenient way to conceive of injective $T_{0}$-spaces is as continuous lattices ${ }^{[12]}$

One of the key properties of continuous lattices are their closure under function spaces, cartesian products and, further, continuity of the standard evaluation and lambda functions:

$$
\begin{aligned}
& \text { ev }:[D \rightarrow D] \times D \rightarrow D \\
& \lambda:[[D \times D] \rightarrow D] \rightarrow[D \rightarrow[D \rightarrow D]]
\end{aligned}
$$

Not only do there exist canonical, continuous evaluation and lambda functions between a continuous lattice $D$ and its corresponding function space, but each continuous lattice can be seen as a projection of its function space.$^{13}$

Definition 5.1. 55] A continuous lattice $(D, \preceq)$ is said to be a projection of a continuous lattice $\left(D^{\prime}, \preceq^{\prime}\right)$ if and only if there is a pair of continuous maps, $i: D \rightarrow D^{\prime}$ and $j: D^{\prime} \rightarrow D$ such that:

$$
j \circ i=i d_{D} \text { and } i \circ j \preceq^{\prime} i d_{D^{\prime}}
$$

In particular, one can observe the following projection of $[D \rightarrow D]$ onto $D$ :

$$
\begin{aligned}
& \Delta: D \rightarrow[D \rightarrow D]: x \mapsto\left(\Delta_{x}: y \mapsto x\right) \\
& \downarrow:[D \rightarrow D] \rightarrow D: f \mapsto f(\perp)
\end{aligned}
$$

Both $\Delta$ and $\downarrow$ are continuous, and $\downarrow$ determines the minimal element in the image of $f$, as continuity of $f$ implies preservation of $\perp$.

One can then consider an $\omega$-sequence of continuous lattices, $\left\langle D_{n}\right\rangle_{0}^{\infty}$, with each $D_{n}$ a projection of $D_{n+1}$, presented by $i_{n}: D_{n} \rightarrow D_{n+1}$ and $j_{n}: D_{n+1} \rightarrow D_{n}$ :

$$
D_{0} \underset{j_{0}}{\stackrel{i_{0}}{\rightleftarrows}} D_{1} \stackrel{i_{1}}{\rightleftarrows} D_{2} \quad \ldots \quad D_{n} \underset{j_{1}}{\stackrel{i_{n}}{\rightleftarrows}} D_{n+1} \quad \ldots
$$

$D_{\infty}$, the inverse limit of the sequence, can then be formed in the standard way, consisting of sequences $\left\langle x_{0}, x_{1}, x_{2}, \ldots\right\rangle$ such that $x_{n} \in D_{n}$ and $j_{n}\left(x_{n+1}\right)=x_{n}$. Interpreting the

[^132]sequence of continuous lattices as a sequence of injective $T_{0}$-spaces, and given the product topology, the limit space inherits the structure of an injective $T_{0}$-space [55, Proposition 4.1]. Furthermore, the canonical projection maps $D_{\infty} \rightarrow D_{n}$ turn out to be projection maps in the sense of Definition 5.1. Thus, at the inverse limit, we obtain a pair of projection maps $j_{\infty_{n}}$ and $i_{n^{\infty}}$ making $D_{n}$ a projection of $D_{\infty}$ for each $n \in N$.

Now, given an arbitrary continuous lattice $D=D_{0}$, we can construct a sequence as above, where $D_{n+1}=\left[D_{n} \rightarrow D_{n}\right]$. An intuitive choice for the $\omega$-sequence of projection pairs $\left\langle i_{n}, j_{n}\right\rangle$ is $\left\langle\Delta_{n}, \downarrow_{n}\right\rangle$. But, as Scott observes, any sequence of projections [ $D_{n} \rightarrow$ $\left.D_{n}\right] \rightarrow D_{n}$ will do. One can then provide an explicit description of the homeomorphism between $D_{\infty}$ and $\left[D_{\infty} \rightarrow D_{\infty}\right.$ ] 55, Theorem 4.4]:

$$
\begin{aligned}
& i_{\infty}(x)=\coprod_{n=0}^{\infty}\left(i_{n} \infty \circ x_{n+1} \circ j_{\infty_{n}}\right) \\
& j_{\infty}(f)=\coprod_{n=0}^{\infty} i_{n+1 \infty}\left(j_{\infty_{n}} \circ f \circ i_{n \infty}\right)
\end{aligned}
$$

In this way, members of $D_{\infty}$ can be interpreted simultaneously as objects, functions, functionals, etc. This description can be formalized by defining application between objects $x, y \in D_{\infty}$ :

$$
x(y)=\coprod_{n=0}^{\infty} i_{n}\left(x_{n+1}\left(y_{n}\right)\right)
$$

## $T_{0}$-Spaces with Countable Bases

The correspondence between continuous lattices and injective $T_{0}$-spaces leads to a "representation theorem" [55]. Injective $T_{0}$-spaces correspond to retracts of cartesian powers of the Sierpinski space (i.e. spaces of the form $\Omega^{X}$ ), which is itself injective. Equally, one can define the "weak" topology of finite character on any powerset $P X$ [42]. The continuous function space $[P X \rightarrow P X]$ consists of morphisms satisfying the condition:

$$
f(x)=\bigcup\left\{f(y) \mid y \in P_{\aleph_{0}}(x)\right\}
$$

We can say that $f$ is a retract if it is idempotent $(f=f \circ f)$. The set of fixed points of $f$ has the structure of a continuous lattice under the $\subseteq$-ordering.

$$
D_{f}=\{x \in P X \mid f(x)=x\}
$$

In fact, every continuous lattice can be obtained (up to isomorphism) in this way [55].

To serve as a foundation for denotational semantics, the $\lambda$-definable objects of a model of untyped $\lambda$-calculus should correspond to computable functions. Scott provides such a model, $P N$ with the topology of finite character, in [57]. ${ }^{14}$ To speak of computability, one wants to be able to speak of recursive enumerability - this is clearly accomplished in $P N$. However, in addition to being an intuitive model, the representation theorem for continuous lattices can be extended to an embedding theorem: Every $T_{0}$-space with a countable basis can be embedded in $P N$ [57, Theorem 1.6]. Thus, $P N$ is not only a convenient domain for denotational semantics, it is a "universal domain" [57].

Besides being a "universal" countable continuous lattice, $P N$ has an additional advantage. As with the (inverse) limit space $D_{\infty}$, members of $P N$ can be interpreted as both objects and functions (iterating as far as one likes), by using the standard coding relation between $P_{\aleph_{0}} N$ and $N$. Here, given a continuous $f: P N \rightarrow P N$ and $x \in P N$, the "projection" functions of the more abstract model in [55] become $\sqrt{15}$

$$
\begin{aligned}
& \operatorname{graph}(f)=\left\{\langle n, m\rangle \mid m \in f\left(e_{n}\right)\right\} \\
& \operatorname{fun}(x)(y)=\left\{m \mid \exists e_{n} \subseteq y \cdot\langle n, m\rangle \in x\right\}
\end{aligned}
$$

Thus, we can list three important properties of $P N$, as a domain:

1. Topology of finite character permits representation theorem for continuous lattices.
2. Coding of the basis by the members of the space (i.e. coding of finite sets of naturals, by naturals themselves).
3. Countable basis.

If one is interested in the third property, look no further than [57]. The first two properties, however, do not depend on the choice of $N$. In theory, a number of sets $A$ could be endowed with a coding of finite sets and, therefore, a coding of the basis of the

[^133]topology of finite character. In this way, one can study continuous (and distributive) functions in a more general context.

## Distributive vs. Continuous Functions

For a given set $X$, we consider two important restrictions of the function space $P X \Rightarrow$ $P X$. These are determined by the distributive and continuous functions, respectively.

Definition 5.2. A function $f$ is said to be distributive if, for each set $X$ :

$$
f^{\prime} X=\bigcup_{x \in X} f^{\prime}\{x\}
$$

Proposition 5.3. 577 For any set $X$, there is a bijective correspondence between binary relations on $X$ and distributive functions $P X \rightarrow P X$.

Definition 5.4. A function $f$ is said to be continuous if, for each set $X$ :

$$
f^{\prime} X=\bigcup_{A \in P_{\wedge_{0}} X} f^{\prime} A
$$

Example 5.5. For a distributive function $f$ :

$$
f^{\bullet}\{x, y\}=f^{\bullet}\{x\} \cup f^{\bullet}\{y\}
$$

For a continuous function $g$ :

$$
g^{\star}\{x\} \cup g^{\star}\{y\} \subseteq g^{\star}\{x, y\}
$$

Hence, the notion of continuity captures the twin ideas of finite character and positive information - a set may encode more information than the sum of its parts.

Lemma 5.6. 57 The following combinators are $\lambda$-abstracts of continuous functions.

$$
\begin{aligned}
& I=\lambda x \cdot x \\
& K=\lambda x y \cdot x \\
& S=\lambda x y z \cdot x(z)(y(z))
\end{aligned}
$$

In NF, it is useful to conceive of the distributive (total) functions as the set of all functions, applied one level down.

Lemma 5.7. For a function $f, j^{\prime} f=\lambda x . f^{\prime \prime} x$ is a distributive function.

Proof. Given a set X, $f^{\prime \prime} X=\left\{f^{6} x \mid x \in X\right\}$, so clearly $f^{"}\{x\}=\left\{f^{6} x\right\}$. Thus,

$$
\bigcup_{x \in X} f^{\prime \prime}\{x\}=\bigcup_{x \in X}\left\{f^{\prime} x\right\}
$$

The latter is just the definition of $f$ " $X$, hence $j(f)$ is distributive.

Therefore, in NF, we can (externally) inject the collection of total functions into the set of distributive functions. ${ }^{166}$

## Application and Abstraction

Given a collection of sets with a surjective pairing function, one can define distributive and continuous application and abstraction operations.

Definition 5.8 (Application). Given two sets $A$ and $B$ (viewed as combinators), we define the following application operations $\sqrt{17}$

$$
\begin{array}{lr}
\operatorname{app}_{\text {dist }}(A, B)=\{x \mid \exists\langle x, b\rangle \in A . b \in B\} & \text { (Distributive Application) } \\
\operatorname{app}_{\text {cont }}(A, B)=\left\{x \mid \exists\langle x, y\rangle \in A . y \in P_{\aleph_{0}}{ }^{`} B\right\} & \\
\text { (Continuous Application) }
\end{array}
$$

Definition 5.9 (Abstraction). Given a function $f$, one can define combinators abstracting the distributive and continuous fragment of $f$, respectively:

$$
\begin{array}{ll}
\lambda_{\text {dist }} x . f[x]=\left\{\langle z, y\rangle \mid z \in f^{\bullet}\{y\}\right\} & \text { (Distributive Abstraction) } \\
\lambda_{\text {cont }} x . f[x]=\left\{\langle z, y\rangle \mid y=\left\{y_{0}, \ldots, y_{n}\right\} \wedge z \in f^{‘} y\right\} & \text { (Continuous Abstraction) }
\end{array}
$$

## Continuous Functions with Multi-Relations

In [58], Scott extends the implementation of continuous functions in $P N$ to arbitrary sets $\mathcal{A}$, where $\mathcal{A}$ codes all finite sequences of its members (i.e. $\mathcal{A}^{*}=\mathcal{A}$ ).

[^134]Definition 5.10. For a set $A, A^{*}$ denotes the collection of finite sequences of elements of $A$.

$$
A^{*}=\bigcup_{n \in N} A^{n}
$$

$A^{0}$ is defined as $\{\rangle\}$, where ' $\rangle$ ' denotes the empty sequence.
Definition 5.11. An $n$-ary relation $S$ corresponds to a subset, $S \subseteq A^{n}$. A multi-relation $M$ is a subset, $M \subseteq A^{*}$.

Definition 5.12. Continuous application of a multi-relation $M$ on a subset $X \subseteq A$ is defined:

$$
M(X)=\left\{y \mid \exists n \exists x_{1}, \ldots, x_{n} \in X .\left\langle y, x_{1}, \ldots, x_{n}\right\rangle \in M\right\}
$$

As a result, we obtain:

$$
M(X)=\cup\left\{M(E) \mid E \in P_{\aleph_{0}} X\right\}
$$

Lemma 5.13. [58] The image function $M(-): P A \rightarrow P A$, induced by a multi-relation $M$ on $A$, is a continuous function (i.e. a function of "finite character").

The following lemma is not stated precisely, but its content is made explicit in both 57 and 58]:

Lemma 5.14. [58] $A$ subset $M \subseteq A^{*}$ is considered finitely complete if $\rangle \in M$, and whenever $\left\langle y, x_{1}, \ldots, x_{j}\right\rangle \in M$ and $\left\{x_{1}, \ldots, x_{j}\right\} \subseteq\left\{y_{1}, \ldots, y_{k}\right\}$, then $\left\langle y, y_{1}, \ldots, y_{k}\right\rangle \in M$. The "finitely complete" subsets of $A^{*}$ are maximal among those representing the same (continuous) image function, and the collection of "finitely complete" multi-relations is in bijection with the continuous functions $P A \rightarrow P A$.

As $S$ and $K$ are $\lambda$-abstracts of continuous functions, one obtains a combinatory algebra for any set $A$ that codes $A^{*}$.

Remark (Extending the Coding of Sequences). The multi-relation model is convenient for NF, as one is able to produce a homogeneous correspondence between sets and sequences, and therefore a homogeneous form of application. The study of such models in a universe of sets, however, motivates the study of a potentially broader class of "coded"
sequences/sets ${ }^{18}$ Indeed, in NF, we are able to generalize the coding of streams ${ }^{19}$ to a broader class of sequences of arbitrary (successor) ordinal length (Proposition 5.50).

But this also changes the underlying topology, formed from the basis of coded sets. Section 5.5 is a step toward studying the impact on the resulting algebra of functions.

## General Ideas Behind Combinatory Algebra

In a basic sense, combinatory algebra consists of a collection $\mathcal{A}$ of objects (combinators) and a pair of operations:

$$
\begin{array}{ll}
\operatorname{app}(-,-): \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A} & \text { (application) } \\
\lambda:(\mathcal{A} \Rightarrow \mathcal{A}) \rightarrow \mathcal{A} & \text { (abstraction) }
\end{array}
$$

Each combinator $A$ corresponds to a function:

$$
\lambda x \cdot \operatorname{app}(A, x): \mathcal{A} \rightarrow \mathcal{A}
$$

A very basic property we would like is the existence of a combinator, corresponding to the "currying" of $\operatorname{app}(-,-)$ :

$$
\begin{equation*}
F u n \equiv \lambda z \cdot \lambda x \cdot \operatorname{app}(z, x): \mathcal{A} \rightarrow(\mathcal{A} \Rightarrow \mathcal{A}) \tag{Fun}
\end{equation*}
$$

In the other direction, one is interested in the sub-collection, $A b s t \subseteq \mathcal{A} \Rightarrow \mathcal{A}$, of functions that can be successfully abstracted.

Definition 5.15. A function $f$ is considered to be successfully abstracted if:

$$
F u n \circ \lambda x . f[x]=f
$$

Equivalently:

$$
\begin{aligned}
\forall A \cdot f^{`} A= & F u n(\lambda x \cdot f[x])(A) \\
& =\operatorname{app}(\operatorname{app}(\lambda z \cdot \lambda x \cdot \operatorname{app}(z, x), \lambda x \cdot f[x]), A)
\end{aligned}
$$

[^135]Remark (Coherent Application and Abstraction). A weaker version of Definition 5.15 could be given, whereby a function $f$ is said to be "successfully abstracted" if and only if:

$$
\exists A \in \mathcal{A} \cdot \operatorname{Fun}(A)=f
$$

In this version, we do not assume the existence of a canonical abstraction of $f$ (i.e. we consider the possibility that $F u n(\lambda x . f[x]) \neq f$, but there exists some $A$ such that Fun $(A)=f$ ). In the distributive and continuous examples, the application and abstraction operations (Definitions 5.9 and 5.8) make the following implication trivial:

$$
(\exists A \in \mathcal{A} \cdot F u n(A)=f) \Longrightarrow \operatorname{Fun}(\lambda x \cdot f[x])=f
$$

We refer to this implication as coherent application and abstraction.
Remark (Coherence and Choice). Thinking along the lines of universal properties and representation, we see that coherence obviates the need for a choice function among the collection of combinators in the fibre $F^{-1}(f)$, for a given function $f \in$ Abst. In this sense, one might also think of coherent abstraction as a form of comprehension. ${ }^{20}$

For the distributive functions of NF, this analogy is even stronger: by Lemma 5.21, one has a unique correspondence between sets and distributive functions. While continuous functions lack this strong form of extensionality, they can be seen as maximal combinators (Scott refers to them as "look-up tables"). So one obtains a unique representative from $\mathrm{Fun}^{-1}(f)$ by way of a closure operation.

In general, the closure of a combinator $A$ is denoted:

$$
1 A \equiv \lambda x . F u n(A)[x]
$$

This canonical (maximal) representation of "extensionally" equivalent combinators gives rise to the Meyer-Scott axiom for $\lambda$-models [41.

$$
\forall A, B \cdot(\forall X \cdot A(X)=B(X)) \Longrightarrow 1 A=1 B
$$

In a formal sense, therefore, $1(-)$ is an inner choice operator 53]. For multi-relation combinators, the inner choice operator, $\lambda X .1 X$, corresponds to finite completion.

[^136]
## Interpreting the Multi-Relation Model

Remark (Well-Defined vs. Over-Defined). Scott's original model is intended for the study of denotational semantics.

$$
n \in x(y) \sim \text { ' } n \text { ' is a value of ' } x \text { ' applied to ' } y \text { ' }
$$

It interprets subsets of $N$ as multiple valued integers.

$$
\begin{aligned}
& x(y)=\emptyset \sim \text { ' } x \text { ' is undefined at ' } y \text { ' } \\
& x(y)=\{n\} \sim \text { ' } x \text { ' is well-defined at ' } y \text { ' } \\
& \exists m, n \in x(y) . m \neq n \sim \text { ' } x \text { ' is over-defined at ' } y \text { ' }
\end{aligned}
$$

In the more general context of $A^{*} \subseteq A$, the distinction between well-defined and overdefined is less clear. In the specific case of $A=N, A \cap P A=\emptyset$. The original interpretation relies on collections (i.e. elements of $P N$ ) being distinct from data (i.e. $N$ ). This disjointness is not guaranteed by a model of the more general structure, defined in [58].

In particular, NF has the closure condition of being a universe. Quite the opposite of being disjoint, $V$ is a fixed point of $P$ :

$$
P V \cap V=V=P V
$$

So we do not have collections coded by, but distinct from, data. Every set $x$ in NF is either undefined (evaluates to $\emptyset$ ) or well-defined at a given $y$, but never over-defined.

Remark (Defaults and Error Terms). Thinking of sets as programs that evaluate inputs of varying size into a register motivates investigation of how programs evaluate empty registers. $x(\emptyset)$ can only evaluate to a non-empty set if $x$ contains "default" values, returned for any input. In the multi-relation model, Scott permits sequences of length one (e.g $\langle x\rangle$ ), without which one could not form the K combinator. However, we also want the empty set to be reserved as an "error" term, as $x(y)=\emptyset$ occurs when $x$ is undefined at $y$. Thus we may wish instead for a theory where $\emptyset$ evaluates to $\emptyset$ always "error" leads to "break." NF turns out to have the ability to model both interpretations, the latter can be referred to as a model of $\lambda$ calculus with effects.

### 5.1.1 Distributive and Continuous Functions in NF

To implement a multi-relation model in NF, one first needs to check that the relevant machinery is stratified and homogeneous.

1. Any definable (i.e. stratified and homogeneous) function in $V \Rightarrow V$ must have a $\lambda$-abstract.
2. Application must be homogeneous.
3. No set, interpreted as a multi-relation, can be the $\lambda$-abstract of an inhomogeneous functional relation ${ }^{21}$

Remark (Extending Local Functions to $V \Rightarrow V$ ). As a matter of convention, we extend every functional relation in NF to a total function, by mapping objects which do not appear in its domain to $\emptyset$. Formally, the graph of $f: x \rightarrow y$ is defined by the pair $\langle g r(f), y\rangle$, consisting of the graph of $f$ and a "tagged" codomain, $y$. We extend $f$ to a total function by dropping its "tagged" codomain and extending $g r(f)$ to a functional relation on $V \times V$ :

$$
\{\langle x, f(x)\rangle \mid x \in \operatorname{dom}(f)\} \cup\{\langle y, \emptyset\rangle \mid y \notin \operatorname{dom}(f)\}
$$

This has no practical impact (the extended functions already existed in $V \Rightarrow V$ ), but reinforces the interpretation of a (total) function $f$ that maps a set $x$ to $\emptyset$, as being 'undefined at $x$.' Thus, a function may be total in the literal sense of having $V$ as a domain, but informally partial if it maps certain objects to the empty set.

As the conditions describing distributivity and continuity are stratified, the collections of distributive and continuous functions form sets in NF.

Definition 5.16. $\mathcal{D}$ and $\mathcal{C}$ denote the sets of distributive and continuous functions, respectively.

[^137]We can define a pair of (inhomogeneous) $\lambda$-operators, mapping elements of $(V \Rightarrow V)$ to their distributive and continuous fragments in $\mathcal{D}$ and $\mathcal{C}$.

Definition 5.17. NF permits both distributive and continuous abstraction.

- The distributive abstraction map $\lambda_{\text {dist }}:(V \Rightarrow V) \rightarrow \mathcal{D}$ is defined by the action:

$$
\lambda_{\text {dist }} f=\left\{\langle y, x\rangle \mid y \in f^{\bullet}\{x\}\right\}
$$

- The continuous abstraction map $\lambda_{\text {cont }}:(V \Rightarrow V) \rightarrow \mathcal{C}$ is defined by the action ${ }^{22}$

$$
\lambda_{\text {cont }} f=\left\{\left\langle y, x_{1}, \ldots, x_{n}\right\rangle \mid y \in f^{\bullet}\left\{x_{1}, \ldots, x_{n}\right\}\right\}
$$

$\lambda_{\text {dist }}$ and $\lambda_{\text {cont }}$ allow us to code functions (as combinators) one type lower than their original graphs, permitting a homogeneous form of evaluation.
$\lambda$-abstraction of multivariate functions requires an implementation of their curried form. In NF, however, the homogeneous version is typically the uncurried form. The relationship between the untyped calculus of functions and the standard (typed) version is indicated by the following diagram. $\cdot{ }^{23}$


The untyped model needs to eliminate the type builders, $\times$ and $\Rightarrow$. Application, resulting from the decoding of sequences to sets, serves as implicit $\times$-elimination. $\lambda$ abstraction serves as a $\Rightarrow$-eliminator, $(V \Rightarrow V) \rightarrow V$.

Example 5.18 (K-combinator). The K-combinator, $\lambda x y . x$, is the curried form of the projection function $\pi_{1}: x \times y \rightarrow x$. While the latter is homogeneous, the former is of "type": $V \Rightarrow(V \Rightarrow V)$. As a result, the standard implementation of $K$ is not possible in NF. Nevertheless, with an appropriate form of application, $K$ can be implemented as a combinator in the multi-relation model. This furnishes an example of how the resulting combinatory algebra alters the functions NF can define as sets. Again, we are

[^138]not introducing new content into NF but new representations of functions that exist as sets in their uncurried form.

Remark (Looking Ahead). The necessary machinery for a relation/multi-relation model is stratified, assuming one has a homogeneous implementation of pairs/sequences. For the distributive case, Quine pairs are sufficient. But the continuous case is trickier. Surjective pairing allows NF to interpret sets as sequences by nesting pairs, but does not provide a canonical interpretation of a given set as a unique finite sequence of sets. By extension, sets carry no canonical interpretation as multi-relations.

### 5.2 A Calculus of Distributive Functions

The interpretation of sets-as-distributive-functions in NF results results in a fairly weak calculus of functions. It does not give a model of full $\lambda$-calculus ${ }^{24}$ On the other hand, each set corresponds, by abstraction, to a unique function - this is the strong extensionality of Lemma $5.48{ }^{25}$

## Homogeneous Application and Abstraction

Implementation of surjective, homogeneous pairs implies distributive application and abstraction are stratified operations:

$$
\begin{align*}
& x(y) \equiv\left\{z \mid \exists a \in x, b \in y \cdot\left\langle\pi_{0}(a), \pi_{1}(a)\right\rangle=\langle z, b\rangle\right\}  \tag{Definition5.8}\\
& \lambda_{\text {dist }} x . f[x]=\left\{\langle z, y\rangle \mid z \in f^{\iota}\{y\}\right\}
\end{align*}
$$

(Definition 5.9)
Where it is obvious, for Section 5.2, we denote $\lambda_{\text {dist }}$-abstraction as simply $\lambda$-abstraction. In other words, given some function $f$ :

$$
\lambda x . f[x] \equiv \lambda_{\text {dist }} x . f[x]=\left\{\langle z, y\rangle \mid z \in f^{\bullet}\{y\}\right\}
$$

[^139]Application is a homogeneous functional relation $V \times V \rightarrow V$. In fact, it is distributive.
Lemma 5.19. $\operatorname{app}(-,-)=-(-): V \times V \rightarrow V$ is distributive in both variables.
Proposition 5.20 ( $\beta$-equivalence). $\beta$-reduction holds for any total distributive function $f$. In other words, we obtain $\beta$-equivalence:

$$
\forall y . f^{\iota} y=\operatorname{app}_{\text {dist }}\left(\lambda_{\text {dist }} x . f[x], y\right) \equiv \lambda_{\text {dist }} x . f[x](y)
$$

Proof. Straightforward.

$$
\begin{aligned}
& \lambda_{\text {dist }} x . f[x]=\left\{\langle z, w\rangle \mid z \in f^{\iota}\{w\}\right\} \\
& \begin{array}{l}
\lambda_{\text {dist }} x . f[z](y) \quad \text { (Application of the } \lambda \text {-abstract) } \\
\quad=\left\{z \mid \exists\langle z, w\rangle \in \lambda_{\text {dist }} x \cdot f[x] . w \in y\right\} \\
\quad=\left\{z \mid \exists w \in y . z \in f^{\iota}\{w\}\right\} \\
\\
=\bigcup_{w \in y} f^{\iota}\{w\}=f^{\iota} y \quad \text { (As } f \text { is distributive) }
\end{array}
\end{aligned}
$$

Combinators are extensional, as objects in the calculus of distributive functions.

Lemma 5.21 ( $\eta$-equivalence). $\eta$-equivalence holds for any set $y$ :

$$
y=\lambda_{\text {dist }} x . y(x)
$$

A corollary of this is strong extensionality of sets as functions:

$$
\forall x, y \cdot x=y \Longleftrightarrow(\forall z \cdot x(z)=y(z))
$$

## Some Basic Combinators

We use nested sequences to bind multiple variables. In this sense, each set corresponds to a unique, pointwise distributive function $V^{n} \rightarrow V$, for each $n$. The empty set is a fixed point of any distributive function, so we cannot form combinators with "default" values.

We can, however, form combinators that return constants for all "defined" inputs. For this, we use the notation '[combinator $]^{-}$' to indicate the distributive fragment of a given combinator.

Example 5.22 (Some Basic Constructions).

$$
\begin{aligned}
& I=\lambda x \cdot x=\{\langle z, z\rangle \mid z \in V\}=\delta_{V} \\
& K^{-}=\lambda x y \cdot x=\{\langle\langle z, c\rangle, z\rangle \mid z, c \in V\}=V \times \pi_{0}{ }^{\bullet} V \\
& \operatorname{app}(-, y)=A_{y}=\lambda x \cdot x y=\{\langle c,\langle c, z\rangle\rangle \mid z \in y\}
\end{aligned}
$$

Remark (Composition vs. Application). The calculus of distributive functions is obviously related to the $\operatorname{Rel}(\mathcal{N})$, the category of relations and relational composition in NF. The distinction being made is between (relational) composition and application. Even within the calculus of distributive functions, the $\lambda$-term, $\lambda a b . a \circ b$, defining composition of combinators is identical to the set-abstract defining composition of relations.

Lemma 5.23 (Composition).

$$
\lambda a b . a \circ b=\lambda a b z . a(b(z))=\{\langle\langle\langle c, y\rangle,\langle d, y\rangle\rangle,\langle c, d\rangle\rangle \mid c, d, y \in V\}
$$

Proof. The proof is meant to acclimate the reader to what is going on.

$$
a(b(z))=\{c \mid \exists d \in b(z) \wedge\langle c, d\rangle \in a\}
$$

In other words:

$$
c \in a(b(z)) \Longleftrightarrow \exists y \in z \cdot\langle d, y\rangle \in b \wedge\langle c, d\rangle \in a
$$

Now $\lambda$-abstraction proceeds as expected:

$$
\begin{aligned}
& \lambda z \cdot a(b(z))=\{\langle c, y\rangle \mid \exists d \cdot\langle d, y\rangle \in b \wedge\langle c, d\rangle \in a\} \\
& \lambda b z \cdot a(b(z))=\{\langle\langle c, y\rangle,\langle d, y\rangle\rangle \mid\langle c, d\rangle \in a\} \\
& \lambda a b z \cdot a(b(z))=\{\langle\langle\langle c, y\rangle,\langle d, y\rangle\rangle,\langle c, d\rangle\rangle \mid c, d, y \in V\}
\end{aligned}
$$

It is clear, by construction, application will simply invert the operation of abstraction. As a result, one can see:

$$
\lambda a b z \cdot a(b(z))(c)(d)=c \circ d
$$

Remark (Programs that Terminate). In our calculus of distributive functions, a program evaluating to $\emptyset$ terminates - "error" leads to "break." As the Quine pairing function codes pairs based on the intersection of elements with $N$, certain sets have finite bounds on successive use of outputs as programs themselves. In other words, for any stream of sets $y_{0}, y_{1}, y_{2}, \ldots ; x\left(y_{1}\right)\left(y_{2}\right) \ldots$ can only evaluate at finitely many steps before entering a fixed "error" state. One can view these as programs that are guaranteed to terminate.

### 5.2.1 Category of Retracts

Each model of untyped $\lambda$-calculus has a corresponding category of retracts 57]. The definition is sufficiently general to develop such a category for the calculus of distributive functions, as well.

The calculus of distributive functions in NF does not form a model of pure $\lambda$-calculus, so we cannot expect the standard properties of a category of retracts (e.g. cartesian closure) to carry over. Nevertheless, the one-to-one relationship between sets and combinators allows one to determine sets of NF that are well-behaved as types. The resulting category of distributive retracts has products and functions spaces but, like $\mathcal{N}$, does not have an evaluation combinator (it isn't distributive). If we embed the category of distributive retracts into the full category of continuous retracts, the distributive product and function space combinators are $\beta$-equivalent to the continuous ones (i.e. they satisfy the universal properties of products and exponentials). Analogous to our earlier use of relative adjoints, we can then use the evaluation combinator, which exists in the category of continuous retracts, to form the exponential adjunction.

Definition 5.24. A retract is a combinator $A$, such that $A \circ A=A$.
Definition 5.25 (Category of Retracts). A category of retracts $\mathcal{R}$ is formed of retracts as objects. Morphisms $f: A \rightarrow B$ are closed $\lambda$-terms $f=\lambda x . B(f(A(x)))$.

Retracts correspond to "types" (or "subtypes" of the universal type), in the sense that they are idempotent. As such, the category of retracts corresponds to a model of typed $\lambda$-calculus, derived from a model of pure $\lambda$-calculus.

From this point, $\mathcal{R}$ will denote the category of retracts for the calculus of distributive functions in NF.

Definition (Objects of $\mathcal{R}$ ). A set $A \in \mathcal{N}$ is a retract if it satisfies:

$$
\lambda x \cdot A(x)=\lambda x \cdot A \circ A(x)=\{\langle z, x\rangle \mid\langle y, x\rangle \in A \wedge\langle z, y\rangle \in A\}
$$

This is equivalent to $A$ satisfying the properties:

$$
\begin{aligned}
& \forall\langle z, x\rangle \in A \cdot \exists y \cdot\langle y, x\rangle \in A \wedge\langle z, y\rangle \in A \\
& (\langle y, x\rangle \in A \wedge\langle z, y\rangle \in A) \Longrightarrow\langle z, x\rangle \in A
\end{aligned}
$$

Alternatively, we can think of ordered pairs as edges of a digraph whose nodes are sets of $\mathcal{N}$ :


Definition (Morphisms of $\mathcal{R}$ ). A morphism $f: A \rightarrow B$ in $\mathcal{R}$ satisfies the property:

$$
f=\lambda x \cdot B(f(A(x)))=\{\langle w, x\rangle \mid \exists y, z \cdot\langle y, x\rangle \in A,\langle z, y\rangle \in f,\langle w, z\rangle \in B\}
$$

Morphisms of $\mathcal{R}$ can be interpreted as sending fixed points of $A$ to fixed points of $B{ }^{26}$ $f=\lambda x \cdot B(f(A(x)))$ is also equivalent to factorization and composition. Once again, we view ordered pairs as edges of a digraph. Each edge is decorated by the set that contains the corresponding ordered pair:


[^140]
## Products and Function Spaces

$\mathcal{R}$ contains combinators that correspond to products and function spaces. ${ }^{[27}$ Rather than a morphism $f: A \rightarrow B$ being a member of $A \Rightarrow B$, we think of $f$ as a term of type $A \Rightarrow B$.

Definition 5.26 (Fixed Points are Terms). If $u$ is a fixed point of the retract $A$ (i.e. $A(u)=u$ ), we write $u: A$ and consider $u$ as being of type $A$.

Definition 5.27 ([57]). The function space and product combinators are defined:

$$
\begin{array}{lr}
A \rightsquigarrow B=\lambda u . B \circ u \circ A & \text { (function space) } \\
A \otimes B=\lambda u \cdot\left\langle A\left(\pi_{1}(u)\right), B\left(\pi_{2}(u)\right)\right\rangle & \text { (product) }
\end{array}
$$

Terms of type $A \otimes B$ are defined by the condition:

$$
u: A \otimes B \Longleftrightarrow \pi_{1}(u): A \wedge \pi_{2}(u): B
$$

Likewise, a combinator $f$ is a morphism $f: A \rightarrow B$ in $\mathcal{R}$ if and only if it is a fixed point of $A \rightsquigarrow B$. We show this explicitly for the distributive retracts of NF.

The function space combinator $A \rightsquigarrow B$ is defined:

$$
A \rightsquigarrow B=\{\langle\langle w, y\rangle,\langle z, x\rangle\rangle \mid\langle x, y\rangle \in A \wedge\langle w, z\rangle \in B\} \quad(A \rightsquigarrow B)
$$

Application of $A \rightsquigarrow B$ to a combinator $f$ results in the set:

$$
(A \rightsquigarrow B)(f)=\{\langle w, y\rangle \mid\langle z, x\rangle \in f \wedge\langle\langle w, y\rangle,\langle z, x\rangle\rangle \in A \rightsquigarrow B\}
$$

Thus, $f$ is a fixed point of $A \rightsquigarrow B$ if and only if $f: A \rightarrow B$ is a morphism in the category of retracts. Thus, in a literal sense, the terms of $A \rightsquigarrow B$ are exactly the morphisms $A \rightarrow B$.

It is apparent that $A \otimes B$ and $A \rightsquigarrow B$ are coherent, as types, in the sense that their terms are exactly what one would expect of function spaces and cartesian products. But

[^141]one also needs to interpret them as objects of a category - i.e. to confirm they satisfy the expected universal properties in $\mathcal{R}$. As the category of distributive retracts is not cartesian closed, we should not expect $A \rightsquigarrow B$ to be an exponential object. We do, however, expect $A \otimes B$ to satisfy the universal property of a binary product. To see this, we need to define a number of other combinators in $\mathcal{R}$.

In 57 ] the pair and sequence combinators for the $\lambda$-calculus are distinct from the pairing function used in the ambient universe, $P N$. In our case, the (Quine) pairing function is also the pairing combinator. By its construction, one can see it is distributive.

Lemma 5.28. The Quine pairing function $\langle-,-\rangle$ defines a combinator, pair, in $\mathcal{R}$.

$$
\langle-,-\rangle=\vartheta_{0} "(-) \cup \vartheta_{1} "(-)
$$

Proof. The combinators corresponding to the three components of the definition are:

$$
\begin{align*}
& \lambda x \cdot \vartheta_{0} "[x]=\left\{\langle z, y\rangle \mid z=\vartheta_{0}[y]\right\}  \tag{‘}\\
& \lambda x \cdot \vartheta_{1} "[x]=\left\{\langle z, y\rangle \mid z=\vartheta_{1}[y]\right\}  \tag{6}\\
& \text { union }=\lambda x y \cdot x \cup y=\{\langle z, y\rangle, x\rangle \mid z=y \vee z=x\}
\end{align*}
$$

(union)
We combine these to define the combinator pair and the projection combinators $f$ st and snd:

$$
\begin{align*}
& \lambda x y .\langle x, y\rangle=\left\{\langle\langle a, b\rangle, c\rangle \mid\langle a, c\rangle \in \lambda x . \vartheta_{0} "[x] \vee\langle a, b\rangle \in \lambda x . \vartheta_{1} "[x]\right\}  \tag{pair}\\
& \left.\begin{array}{l}
\lambda u . \pi_{1}[u]=\{\langle y, x\rangle \mid 0 \notin x
\end{array}\right) \quad(z \in(y \cap N) \Leftrightarrow z+1 \in(x \cap N))  \tag{fst}\\
& \wedge(z \in(y-N) \Leftrightarrow z \in(x-N))\} \\
& \lambda u . \pi_{2}[u]=\{\langle y, x\rangle \mid 0 \in x \wedge(z \in(y \cap N) \Leftrightarrow z+1 \in(x \cap N))  \tag{snd}\\
& \wedge
\end{align*}
$$

Just as we have defined the pairing combinator, we can define the special case corresponding to the diagonal combinator:

$$
\begin{equation*}
\lambda u .\langle u, u\rangle=\left\{\langle x, y\rangle \mid\langle x, y\rangle \in \lambda x \cdot \vartheta_{0} "[x] \vee\langle x, y\rangle \in \lambda x \cdot \vartheta_{1} "[x]\right\} \tag{diag}
\end{equation*}
$$

Lemma 5.29. $u=\left\langle\pi_{1}(u), \pi_{2}(u)\right\rangle$ or, in terms of the combinators defined above:

$$
\lambda u \cdot p a i r(f s t(u))(\operatorname{snd}(u))=I
$$

Proof. The set that realizes $\lambda u$.pair $(f \operatorname{st}(u))(\operatorname{snd}(u))$ is:

$$
\left\{\langle a, x\rangle \mid\left(\langle a, c\rangle \in \lambda x \cdot \vartheta_{0} "[x] \wedge c \in f s t(\{x\})\right) \vee\left(\langle a, c\rangle \in \lambda x \cdot \vartheta_{1} "[x] \wedge c \in \operatorname{snd}(\{x\})\right)\right\}
$$

As a singleton set determines a pair of the form $\langle y, \emptyset\rangle$ or $\langle\emptyset, y\rangle$, and $\vartheta$ is inverse to $\pi$, $a=x$ for all $\langle a, x\rangle$ in the set. Therefore:

$$
\lambda u \cdot \operatorname{pair}(f \operatorname{st}(u))(\operatorname{snd}(u))=\delta_{V}=I
$$

Definition (NF). The set abstract corresponding to the product combinator $A \otimes B$ is defined:

$$
\begin{align*}
& \lambda u .\left\langle A\left(\pi_{1}(u)\right), B\left(\pi_{2}(u)\right)\right\rangle \\
& \quad=\lambda u \cdot p a i r(A(f \operatorname{st}(u)))(B(\operatorname{snd}(u))) \\
& =
\end{align*} \quad\left\{\langle y, x\rangle \mid\left(\exists\langle z, x\rangle \in f s t .\langle w, z\rangle \in A \wedge\langle y, w\rangle \in \lambda u . \vartheta_{0} "[u]\right)\right\}
$$

Thus $A \otimes B$ corresponds to the completion "across the middle" of the diagram below, where either the top or the bottom three arrows exist (recall, ordered pairs are coded by $\vartheta_{0} \cup \vartheta_{1}$ ):


We can now see, explicitly, for the category of distributive NF-retracts:

$$
\begin{aligned}
& f \text { st } \circ(A \otimes B):(A \otimes B) \rightsquigarrow A \\
& \text { snd } \circ(A \otimes B):(A \otimes B) \rightsquigarrow A
\end{aligned}
$$

This allows us to carry out the standard proof of the universal property.
Proposition 5.30. The category $\mathcal{R}$ of retracts of the calculus of distributive functions of NF has products and function spaces (though not necessarily cartesian closure). Furthermore, it has a universe object, $V$.

Remark (Weak vs. Strong Universe Objects). The content of Proposition 5.30 indicates the similarity between $\mathcal{R}$ and $\mathcal{N}$. Clearly, $V \circ V=V$. So the universal set is a (distributive) retract and, therefore, the $\subseteq$-maximal retract. However, the universe object of $\mathcal{R}$ is not all that we might have hoped it would be. The universal retract we want is far more in the nature of a type-classifier, and the terms of $V$ are hardly equivalent to the objects of $\mathcal{R}{ }^{28}$ Rather than:

$$
x: V \Longleftrightarrow x \circ x=x
$$

We obtain only:

$$
x: V \Longleftrightarrow x=V \vee x=\emptyset
$$

Remark (Cartesian Closure in $\mathcal{R}$ ). Cartesian closure in $\mathcal{R}$ mimics the unit and counit of the corresponding adjoint pair, via the combinators curry and eval. Even in the case of distributive NF-retracts, we can form curry, reflecting the homogeneity of application in the calculus of distributive functions.

$$
\begin{equation*}
\lambda u x y \cdot u(\langle x, y\rangle)=\{\langle\langle\langle c, y\rangle, x\rangle,\langle c, b\rangle\rangle \mid\langle\langle b, y\rangle, x\rangle \in \operatorname{pair}\} \tag{curry}
\end{equation*}
$$

Evaluation corresponds to the combinator:

$$
e v a l=\lambda x \cdot \pi_{1}(x)\left(\pi_{2}(x)\right)
$$

Cartesian closure of $\mathcal{R}$ would require eval $\circ((A \rightsquigarrow B) \otimes A)$ to be a term of type $((A \rightsquigarrow B) \otimes A) \rightsquigarrow B$.

$$
\text { eval } \circ((A \rightsquigarrow B) \otimes A):((A \rightsquigarrow B) \otimes A) \rightsquigarrow B
$$

But eval is not distributive, so does not correspond to an object of $\mathcal{R}$.

[^142]We do have the combinator $\lambda x y \cdot \pi_{1}(x)\left(\pi_{2}(y)\right)$, so given some $u:(A \rightsquigarrow B) \otimes A$ we could form:

$$
\lambda x y \cdot \pi_{1}(x)\left(\pi_{2}(y)\right)(u)(u)=\lambda x \cdot \pi_{1}(x)\left(\pi_{2}(x)\right)(u) \sim \operatorname{eval}(u)
$$

But this is not the same as defining a universal transformation ${ }^{29}$

### 5.3 A Model of The $\lambda$-Calculus

Continuous functions require the implementation of sequences $\left\langle c, x_{0}, \ldots, x_{n}\right\rangle$, where $c \in$ $\tau^{`}\left\{x_{0}, \ldots x_{n}\right\}$. Iteration of Quine pairing allows one to code a given finite sequence as a set in NF. Furthermore, given any length $n$, the implementation results in an exact equivalence $V^{n}=V$. But this is not sufficient to form a model of untyped $\lambda$-calculus.
$\lambda$-abstraction requires the implementation of sequences of varying finite length. For a combinator $X$ to be $\beta$-equivalent to a continuous function, application needs to associate members of $X$ to sequences of varying length. While standard iteration of Quine pairing can code a given sequence as a set, sets themselves do not code canonical finite sequences, unless a fixed length is imposed externally. In other words, iterated Quine pairing allows one to form code $V^{n}$, for any given $n$, but does not permit the formation of a "cumulative" collection of sequences:

$$
V^{*}=\bigcup_{n \in N} V^{n}
$$

The primary objective of this section is to develop a means of associating a canonical finite sequence with each set, in such a way that $V^{*} \subseteq V$, while attempting to minimize redundancy (i.e. multiple sets coding the same sequence). From this we obtain a multirelation model of pure $\lambda$-calculus.

[^143]
### 5.3.1 Coding Sequences of Varying Length

Surjectivity of Quine pairs allows us to consider the following partition of $V$ :

$$
V=(N \times V) \cup(V \backslash N \times V)
$$

We associate each set $z$ with a unique finite sequence by considering its corresponding Quine pair $\left\langle\pi_{1}(z), \pi_{2}(z)\right\rangle$.

Definition 5.31 (Pairs as Finite Sequences). The interpretation of a pair $\langle c, x\rangle$ as a finite sequence has two possible cases, depending on where $\langle c, x\rangle$ falls in the partition of $V$ :

$$
\begin{array}{lr}
\langle c, x\rangle \sim\langle c, x\rangle & (\langle c, x\rangle \in V \backslash N \times V) \\
\langle c,\langle w, y\rangle\rangle \sim\left\langle w, y_{1}, \ldots, y_{c}\right\rangle & (\langle c, x\rangle \in N \times V)
\end{array}
$$

Where $x=\langle w, y\rangle$ and $\left\langle y_{1}, \ldots, y_{c}\right\rangle$ is the $c$-length sequence corresponding to $y$.
A set of the form $z=\langle 0, c\rangle$ is interpreted as the single element sequence: $\langle c\rangle$. The empty sequence, $\rangle$, is formed by $\emptyset$.

The coding of sets-as-sequences, given in Definition 5.31 achieves the goal of associating a unique finite sequence to each set, in such a way that $V^{*} \subset V$. But there is redundancy in the representation of sequences-as-sets. For $c \notin N,\langle 1,\langle c, x\rangle\rangle$ and $\langle c, x\rangle$ correspond to the same ordered pair.

But this redundancy only occurs on the left side of the application operation. It may be the case that distinct sets $a$ and $b$ satisfy the property:

$$
\forall x \cdot a(x)=b(x)
$$

but, as arguments, distinct $a$ and $b$ must satisfy:

$$
\neg \forall x . x(a)=x(b)
$$

### 5.3.2 Implementation of Application and Abstraction

Definition 5.31 implies the following definitions of application and abstraction.
Definition 5.32 (Application). Application of two sets $a$ and $b$ is defined:

$$
\begin{aligned}
a(b)=\{c \mid(c \notin N & \wedge \exists x \in b .\langle c, x\rangle \in a) \\
& \left.\vee\left(\exists n \in N .\langle n,\langle c, x\rangle\rangle \in a .\left\{x_{1}, \ldots, x_{n}\right\} \subset b\right)\right\}
\end{aligned}
$$

Definition 5.33 (Abstraction). Given a continuous total function $\tau$, the $\lambda$-abstraction of $\tau$ is defined:

$$
\begin{aligned}
\lambda x . \tau[x]= & \left\{\langle c, x\rangle \mid c \notin N \wedge c \in \tau^{\star}\{x\}\right\} \\
& \cup\left\{\langle n,\langle c, x\rangle\rangle \mid n \in N \wedge c \in \tau^{‘}\left\{x_{1}, \ldots x_{n}\right\}\right\}
\end{aligned}
$$

Proceeding further requires a more formal analysis, to ensure the relevant formulas are stratified. It is helpful to think of Definition 5.31 as coding two pieces of information: a single element (of the same type as the pair) and a finite set (one type above the pair). We refer to these coding operations as $\alpha$ and $\chi$, respectively.

Definition $5.34(\alpha)$. Given any set $z$ of NF, we interpret it as an ordered triple $\left\langle z_{1}, z_{2}, z_{3}\right\rangle$. The graph of $\alpha: V \rightarrow V$ is defined by partitioning the universe:

$$
\alpha \equiv \pi_{1} \upharpoonright(V \backslash N \times V \times V) \cup \pi_{2} \upharpoonright(N \times V \times V)
$$

The action of $\alpha$ is defined:

$$
\begin{array}{lr}
\alpha:\langle c, y\rangle \mapsto c & (V \backslash N \times V \times V) \\
\alpha:\langle n,\langle c, x\rangle\rangle \mapsto c & (N \times V \times V)
\end{array}
$$

The definition of $\chi$ is less straightforward, as it is inhomogeneous. Taking the partition used in Definition 5.31, we define $\chi$ in two parts:

$$
\chi=\chi \upharpoonright(V \backslash N \times V) \cup \chi \upharpoonright(N \times V)
$$

The first case, $\langle c, x\rangle \in V \backslash N \times V$, is straightforward.

Definition $5.35(\chi \upharpoonright V \backslash N \times V)$. A stratified, type-raising operation $\chi \upharpoonright V \backslash N \times V$ is defined by the action:

$$
\chi \upharpoonright(V \backslash N \times V):\langle c, x\rangle \mapsto\{x\}
$$

In the second case, we interpret $\langle n, y\rangle \in N \times V$ as $\langle n,\langle c, x\rangle\rangle$, where $\langle c, x\rangle=y$. We need to formalize the operation:

$$
\chi:\langle n,\langle c, x\rangle\rangle \mapsto\left\{x_{1}, \ldots, x_{n}\right\} \equiv\left\{w \mid \exists i \leq n \cdot x_{i}=w\right\}
$$

Formally, the use of ' $x_{i}=w$ ' implies the existence of a homogeneous projection function $\pi: N \times V \rightarrow V$, where $x_{i}=\pi(i, x) .30$ However, $\pi$ is not independent of the length assigned to the sequence derived from $x{ }^{31}$ We need to define an operation $\bar{\pi}$, associating each $n \in N$ with a homogeneous projection function $\pi_{n}: N \times V \rightarrow V$ :

$$
\bar{\pi}: N \rightarrow(N \times V \Rightarrow V) ; n \mapsto \pi_{n}(-,-): N \times V \rightarrow V
$$

The association of $n$ and $x$ with a sequence $\left\langle x_{1}, \ldots, x_{n}\right\rangle$ can be formalized by taking $\bar{\pi}$ in its un-curried (homogeneous) form:

$$
\bar{\pi}: N \times N \times V \rightarrow V
$$

Each member $x_{i}$ of $\left\langle x_{1}, \ldots, x_{n}\right\rangle$ corresponds $\pi_{n}(i, x)$, where:

$$
\pi_{n}(-, x) \equiv \bar{\pi}(n,-, x): N \rightarrow V
$$

Technically, we are identifying a stream with an $n$-length sequence:

$$
\left\langle x_{1}, \ldots, x_{n}, \emptyset, \emptyset, \ldots\right\rangle \sim\left\langle x_{1}, \ldots, x_{n}\right\rangle
$$

But this is not an issue, as the set abstract we are interested in formalizing only uses the restriction of $\bar{\pi}(n,-, x)$ to $\{1, \ldots, n\}{ }^{32}$

[^144]Definition $5.36(\chi \upharpoonright N \times V)$. Using the identity $N \times V=N \times V \times V, \chi \upharpoonright N \times V$ is a stratified, type-raising operation:

$$
\chi \upharpoonright N \times V \times V: z \mapsto\{w \mid \exists n \leq \bar{\pi}(3,1, z) \cdot \bar{\pi}(\bar{\pi}(3,1, z), n, \bar{\pi}(3,3, z))=w\}
$$

We can now replace the informal notation in Definitions 5.32 and 5.33 with the formal, stratified operations $\alpha$ (Definition 5.34) and $\chi$ (Definitions 5.35 and 5.36), which are homogeneous and type-raising, respectively:

$$
\begin{array}{ll}
A(B)=\{c \mid \exists a \in A . \alpha(a)=c \wedge \chi(a) \subseteq B\} \\
\lambda x \cdot \tau[x]=\{z \mid \alpha(z) \in \tau \circ \chi(z)\} & \text { (Application) }
\end{array}
$$

From the stratified operations defined above, we obtain an important lemma.
Lemma 5.37. The application operation, given in Definition 5.32, is stratified and homogeneous. Thus, we can implement application and abstraction of multi-relations in any model of $N F$.

Unrestricted (stratified) comprehension implies a further result.
Lemma 5.38. $\mathcal{D} \subsetneq \mathcal{C} \subsetneq(V \Rightarrow V)$

Proof. The function $x \mapsto x(x)$ is continuous but not distributive. The complementation function $x \mapsto V \backslash x$ is anti-continuous.
$\lambda^{\prime \prime} \mathcal{C}=\{\lambda x . \tau[x] \mid \tau \in \mathcal{C}\}$ is a type below $\mathcal{C}$, but we can still prove $\beta$-equivalence.
Lemma 5.39 ( $\beta$-Equivalence). For a continuous function $\tau$,

$$
\forall Y \cdot \lambda x \cdot \tau[x](Y)=\tau^{\prime} Y
$$

Proof. $\subseteq$ Suppose $c \in \lambda x \cdot \tau[x](Y)$. Then, by definition of the $\lambda$-abstraction:

$$
\exists y \in Y . c \in \tau^{\star}\{y\} \vee \exists\left\{y_{1}, \ldots, y_{n}\right\} \subseteq Y . c \in \tau^{\star}\left\{y_{1}, \ldots, y_{n}\right\}
$$

Hence

$$
c \in \cup\left\{\tau^{\prime} z \mid z \in P_{\aleph_{0}}(Y)\right\}
$$

As $\tau$ is continuous, $c \in \tau^{\iota} Y$.
$\supseteq$ Suppose $c \in \tau^{\ell} Y$. As $\tau$ is continuous, there is some finite subset $z$ of $Y$, such that $c \in \tau^{\prime} z$. Therefore, by definition, $c \in \lambda x \cdot \tau[x](Y)$.

The $\xi^{*}$-rule (as defined in [57]) also holds, for any $\sigma, \tau \in \mathcal{C}$ :

$$
\begin{equation*}
\lambda x . \tau \subseteq \lambda x . \sigma \leftrightarrow \forall x . \tau^{6} x \subseteq \sigma^{6} x \tag{*}
\end{equation*}
$$

### 5.3.3 Implementation of Combinators

The results summarized in Lemma 5.37 prove that we can implement a multi-relation model in any model of NF. We now discuss some of the set-abstracts that correspond to combinators of particular interest: $S, K, I$ and, the fixed point combinator, $Y(I)$.

Definition $5.40(I$ and $K) .{ }^{33}$

$$
\begin{aligned}
I= & \{\langle x, x\rangle \mid x \notin N\} \cup\left\{\langle n,\langle c, x\rangle\rangle \mid c \in\left\{x_{1}, \ldots, x_{n}\right\}\right\} \\
K= & \left\{\langle n,\langle\langle m,\langle c, z\rangle\rangle, x\rangle\rangle \mid c \in\left\{x_{1}, \ldots, x_{n}\right\}\right\} \\
& \cup\left\{\langle n,\langle\langle c, z\rangle, x\rangle\rangle \mid c \notin N \wedge c \in\left\{x_{1}, \ldots, x_{n}\right\}\right\} \\
& \cup\{\langle\langle c, z\rangle, x\rangle \mid c \notin N \wedge c=x\} \\
& \cup\{\langle\langle n,\langle c, z\rangle\rangle, x\rangle \mid c=x\}
\end{aligned}
$$

$\lambda$-abstracts, formulated as "look-up tables," correspond to "maximal" multi-relations [57. As a result, the set abstract of $K$ has already spiraled into 4 cases (if we bind $n$ variables we get $2^{n}$ cases).

To preserve even a weak form of $\eta$-equivalence:

$$
\forall y . y \subseteq \lambda x . y(x)
$$

[^145]we need to maintain "look-up tables" as the formal means of abstraction. Practically speaking, however, $K$ requires only the first of the four abstracts. Redundancy in the coding of sequences permits elimination of the remaining elements in the look-up table.

Lemma 5.41. $\forall x \cdot K^{*}(x)=K(x)$, where

$$
K^{*}=\left\{\langle n,\langle\langle m,\langle c, z\rangle\rangle, x\rangle\rangle \mid c \in\left\{x_{0}, \ldots, x_{n}\right\}\right\}
$$

Proof. Given any two sets, $a$ and $b$, we need to prove $K^{*}(a)(b)=a$.

$$
K^{*}(a)=\left\{\langle m,\langle c, z\rangle\rangle \mid \exists\langle n,\langle\langle m,\langle c, z\rangle\rangle, x\rangle\rangle \in K^{*} \wedge\left\{x_{0}, \ldots, x_{n}\right\} \subseteq a\right\}
$$

So $K^{*}(a)$ returns the set of all $\langle m,\langle c, z\rangle\rangle$, where $c \in a$. In other words, $K^{*}(a)$ codes $a \times B_{\text {Fin }}$ " $a{ }^{34}$ As any set $b$ is "covered" by its finite subsets, $K^{*}(a)(b)=a$.

This is a specific case of a more general phenomenon. We frequently invoke [combinator] ${ }^{*}$, the abbreviated version of [combinator], where the full version is easily obtained by undoing the identification $\langle c, x\rangle \sim\langle 1,\langle c, x\rangle\rangle$ at each step.

Definition 5.42 (Abstraction*). Given a continuous total function $\tau$, the minimal $\lambda$ abstract of $\tau$ is defined:

$$
\lambda x \cdot \tau[x]^{*}=\left\{\langle n,\langle c, x\rangle\rangle \mid n \in N \wedge c \in \tau^{\star}\left\{x_{0}, \ldots x_{n}\right\}\right\}
$$

It is important to note, however, we must not identify [combinator] and [combinator]* when they occur as arguments, on the right-hand side of application. In general:

$$
Z(\lambda x \cdot \tau[x]) \neq Z\left(\lambda x . \tau[x]^{*}\right)
$$

Remark (Binding and Nesting). In Lemma 5.41, we defined $\lambda x y . x$ as the set:

$$
K^{*}=\left\{\langle n,\langle\langle m,\langle c, z\rangle\rangle, x\rangle\rangle \mid c \in\left\{x_{0}, \ldots, x_{n}\right\}\right\}
$$

The members of $K^{*}$, interpreted as a multi-relation, correspond to a nested pair of finite sequences that, in turn, correspond to the binding of a pair of variables:

$$
\langle n,\langle\langle m,\langle c, z\rangle\rangle, x\rangle\rangle \sim\left\langle\left\langle c, z_{0}, \ldots, z_{m}\right\rangle, x_{0}, \ldots, x_{n}\right\rangle
$$

[^146]In general, the binding of an $n$-variable continuous function (by $\lambda$-abstraction) results in a set, each of whose members is a nesting of $n$ finite sequences.

Remark (Presentation of More Complex Combinators). Proving that any model of NF has an interpretation as a (multi-relation) model of $\lambda$-calculus in NF has, in many ways, already been achieved. With the development of a method for implementing iterative/nested finite sequences (Definition 5.31), in a way that satisfies $V^{*} \subseteq V$, and application/abstraction operations satisfying Lemma 5.37, the remainder is, in effect, corollary to [58] ${ }^{35}$ But our interest in the model extends beyond its existence.

The multi-relation model untyped $\lambda$-calculus, in NF, not only satisfies the general condition given in [58], but satisfies a particularly strong case where the ambient set is a fixed point of $P$ (i.e. $P V=V$ ). So we are also interested in the semantics. As such, for the remainder of this section, we are opting for a middle path. To prove the existence of a given combinator in our model, we state the set abstracts explicitly. For the proofs, we use the generic embedded sequences coded by members of our set to define the corresponding abstract multi-relation and refer the reader to [58].

Lemma 5.43 (Composition). The following set abstract defines the composition combinator (displayed in its reduced form, comp*):

$$
\begin{aligned}
& c o m p^{*}=a \circ b=\lambda a b x \cdot a(b(x)) \\
& \qquad=\left\{\langle k,\langle\langle j,\langle\langle n,\langle c, x\rangle\rangle, s\rangle\rangle, t\rangle\rangle \mid \exists\langle m,\langle c, z\rangle\rangle \in\left\{t_{0}, \ldots, t_{k}\right\} .\right. \\
& \left.\quad \forall z_{i} \exists\left\langle m_{i},\left\langle z_{i}, y^{i}\right\rangle\right\rangle \in\left\{s_{0}, \ldots, s_{j}\right\} .\left\{y_{0}^{i}, \ldots, y_{m_{i}}^{i}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}\right\}
\end{aligned}
$$

Proof. Consider the set to which we would like $\operatorname{comp}(a)(b)(w)$ to reduce.

$$
a(b(w))=\left\{c \mid\left\langle c, x_{0}, \ldots, x_{n}\right\rangle \in a \wedge \forall x_{i} \exists\left\langle x_{i}, z_{0}, \ldots z_{n_{i}}\right\rangle \in b .\left\{z_{0}, \ldots, z_{n_{i}}\right\} \subseteq w\right\}
$$

We bind the variables by restricting them to generic finite sets, in the formula defining

[^147]$\lambda$-abstract, $\lambda x \cdot a(b(x))$.
\[

$$
\begin{aligned}
& \left\langle\left\langle\left\langle c, x_{0}, \ldots, x_{n}\right\rangle, s_{0}, \ldots, s_{j}\right\rangle, t_{0}, \ldots t_{k}\right\rangle \in \lambda x \cdot a(b(x)) \\
& \Longleftrightarrow
\end{aligned}
$$
\]

$$
\exists\left\langle c, z_{0}, \ldots, z_{m}\right\rangle \in\left\{t_{0}, \ldots, t_{k}\right\} . \forall z_{i} \exists\left\langle z_{i}, y_{0}, \ldots, y_{m_{i}}\right\rangle \in\left\{s_{0}, \ldots, s_{j}\right\} \cdot\left\{y_{0}, \ldots, y_{m_{i}}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}
$$

Proving $\beta$-equivalence is straightforward.

Theorem 5.44 (The $S$ Combinator). The following set abstract defines the $S$ Combinator (displayed in its reduced form, $S^{*}$ ):

$$
\begin{aligned}
& S^{*}= \lambda a b u \cdot a(u)(b(u)) \\
&=\{\langle k,\langle\langle j, \\
&\quad \exists\langle n,\langle c, x\rangle\rangle, s\rangle\rangle, t\rangle\rangle \mid \\
& \exists\left\langle n^{\prime},\right.\left.\left\langle\langle m,\langle c, z\rangle\rangle, x^{\prime}\right\rangle\right\rangle \in\left\{t_{0}, \ldots, t_{k}\right\} \cdot\left\{x_{0}^{\prime}, \ldots, x_{n^{\prime}}^{\prime}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\} \\
&\left.\wedge \forall z_{i} \cdot \exists\left\langle n_{i},\left\langle z_{i}, y^{i}\right\rangle\right\rangle \in\left\{s_{0}, \ldots, s_{j}\right\} \cdot\left\{y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}\right\}
\end{aligned}
$$

Proof.

$$
\left.\begin{array}{rl}
a(u)(b(u))=\left\{c \mid \exists\left\langle\left\langle c, z_{0}, \ldots, z_{m}\right\rangle, x_{0}^{\prime}, \ldots,\right.\right. & \left.x_{n^{\prime}}^{\prime}\right\rangle
\end{array} \quad \in a \cdot\left\{x_{0}^{\prime}, \ldots, x_{n^{\prime}}^{\prime}\right\} \subseteq u\right\}
$$

Binding the variables (' $u$,' in particular) requires

$$
\left\{x_{0}^{\prime}, \ldots, x_{n^{\prime}}^{\prime}\right\} \cup\left[\bigcup_{z_{i} \in\left\{z_{0}, \ldots, z_{m}\right\}}\left\{y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\}\right]
$$

to be a finite set. It obviously is, but notice this implies a closure property on our abstract coding operation $\chi \underbrace{36}$ The appropriate definition of $S$ is given by:

$$
\begin{aligned}
&\left\langle\left\langle\left\langle c, x_{0}, \ldots, x_{n}\right\rangle, s_{0}, \ldots, s_{j}\right\rangle, t_{0}, \ldots t_{k}\right\rangle \in \lambda x y z . x(z)(y(z)) \\
& \Longleftrightarrow \\
& \quad \exists\left\langle\left\langle c, z_{0}, \ldots, z_{m}\right\rangle, x_{0}^{\prime}, \ldots, x_{n^{\prime}}^{\prime}\right\rangle \in\left\{t_{0}, \ldots, t_{k}\right\} \cdot\left\{x_{0}^{\prime}, \ldots, x_{n^{\prime}}^{\prime}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\} \\
& \wedge \forall z_{i} \exists\left\langle z_{i}, y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\rangle \in\left\{s_{0}, \ldots, s_{j}\right\} \cdot\left\{y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}
\end{aligned}
$$

[^148]Remark (The $S$-Combinator and Choice). The proof of Theorem 5.44 noted:

$$
c \in a(u)(b(u)) \Longleftrightarrow \exists\left\langle c, z_{0}, \ldots, z_{m}\right\rangle \in a(u) \cdot\left\{z_{0}, \ldots, z_{m}\right\} \subseteq b(u)
$$

Written more formally, the subset condition is:

$$
\forall z_{i} \in\left\{z_{0}, \ldots, z_{m}\right\} \cdot \exists\left\langle z_{i}, y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\rangle \in b . \forall y_{j} \in\left\{y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\} \cdot y_{j} \in u
$$

But the existential quantifier does not, in general, imply the existence of a canonical sequence in $b$, for each $z_{i}$. While $\left\{z_{0}, \ldots, z_{m}\right\}$ is finite, the following set could easily be infinite:

$$
\left\{\langle w, \vec{y}\rangle \in b \mid w \in\left\{z_{0}, \ldots, z_{m}\right\}\right\}
$$

Thus, our formation of the sum-set:

$$
\bigcup_{z_{i} \in\left\{z_{0}, \ldots, z_{m}\right\}}\left\{y_{0}^{i}, \ldots, y_{n_{i}}^{i}\right\}
$$

requires choosing a unique sequence from each member of the $\left\{z_{0}, \ldots z_{m}\right\}$-indexed family of (pairwise disjoint) sets:

$$
\coprod_{z_{i} \in\left\{z_{0}, \ldots, z_{m}\right\}}\left\{\langle w, \vec{y}\rangle \in b \mid w=z_{i}\right\}
$$

Of course, this is a finite (i.e. trivial) choice function. It is worth noting, however, to consider more general forms of continuity, one may require a nontrivial form of choice. Therefore, if one interprets the basis of a more general topology as consisting of the "small" sets, the property discussed in Section 3.6, the sum-set of a small family of small sets is small, may not be sufficient for the formation of a continuous model of combinatory algebra. ${ }^{[77}$

Theorem 5.45. (Summary) Any model of NF gives a model of the lambda calculus, with $S, K$, and I defined as above.

[^149]
## The Fixed Point Combinator

One aspect of semantic interest, regarding our model, is the set in NF that corresponds to the paradoxical combinator applied to identity combinator:

$$
\begin{aligned}
& Y=\lambda f \cdot \lambda x \cdot f(x x)(\lambda x \cdot f(x x)) \\
& Y(I)=\lambda x \cdot x x(\lambda x \cdot x x)
\end{aligned}
$$

Determining the set that corresponds to $Y$ requires the construction of a further pair of combinators:

$$
\begin{aligned}
& \lambda x \cdot x x= \\
& \quad\left\{\langle n,\langle c, x\rangle\rangle \mid \exists\left\langle n^{\prime},\left\langle c, x^{\prime}\right\rangle\right\rangle \in\left\{x_{0}, \ldots, x_{n}\right\} \cdot\left\{x_{0}^{\prime}, \ldots x_{n^{\prime}}^{\prime}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}\right\} \\
& \lambda x \cdot a(a(x))=\quad \text { (Self Application) } \\
& \quad\left\{\langle n,\langle c, x\rangle\rangle \mid \exists\left\langle n^{\prime},\left\langle c, x^{\prime}\right\rangle\right\rangle \in a . \forall x_{i}^{\prime} \exists\left\langle n_{i},\left\langle x_{i}^{\prime}, z^{i}\right\rangle\right\rangle \in a .\left\{z_{0}^{i}, \ldots, z_{n_{i}}^{i}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\}\right\}
\end{aligned}
$$

The fixed point of the identity combinator turns out to be equal to the universe:

$$
Y(I)=V
$$

Theorem 5.46. $\lambda x \cdot x x(\lambda x . x x)=V$

Proof. As we showed above, $\langle n,\langle c, x\rangle\rangle \in \lambda x . x x$ exactly when there is an element of $\left\{x_{0}, \ldots, x_{n}\right\}$ coding a subset of $\left\{x_{0}, \ldots, x_{n}\right\}$ that "evaluates" to $c$.

We fully unpack the set corresponding to $\lambda x \cdot x x(\lambda x . x x)$ :

$$
\begin{aligned}
\lambda x . x x & (\lambda x . x x) \\
& =\left\{c \mid \exists\langle n,\langle c, x\rangle\rangle \in \lambda x . x x .\left\{x_{0}, \ldots, x_{n}\right\} \subseteq \lambda x . x x\right\} \\
= & \left\{c \mid \exists\langle n,\langle c, x\rangle\rangle \cdot \exists\left\langle n^{\prime},\left\langle c, x^{\prime}\right\rangle\right\rangle \in\left\{x_{0}, \ldots, x_{n}\right\} .\right. \\
& \left.\quad\left\{x_{0}^{\prime}, \ldots x_{n^{\prime}}^{\prime}\right\} \subseteq\left\{x_{0}, \ldots, x_{n}\right\} \subseteq \lambda x . x x\right\}
\end{aligned}
$$

Thus, $\left\langle n^{\prime},\left\langle c, x^{\prime}\right\rangle\right\rangle \in \lambda x$.xx, so there must be some $\langle m,\langle c, z\rangle\rangle \in\left\{x_{0}^{\prime}, \ldots x_{n^{\prime}}^{\prime}\right\}$ such that $\left\{z_{0}, \ldots, z_{m}\right\} \subseteq\left\{x_{0}^{\prime}, \ldots x_{n^{\prime}}^{\prime}\right\}$. We quickly see that this forms an infinite descending sequence
of finite sets, unless there is a fixed point:

$$
\langle m,\langle c, z\rangle\rangle \in\left\{z_{0}, \ldots, z_{m}\right\} \subseteq \lambda x . x x
$$

As our initial choice $\left\{x_{0}, \ldots, x_{n}\right\}$ was finite, we cannot have an infinite descending sequence of subsets. We must have a least element $\langle m,\langle c, z\rangle\rangle$ that will be a fixed point.

So our proof reduces to showing:

$$
c \in \lambda x \cdot x x(\lambda x . x x) \Longleftrightarrow \exists m \in N, z .\langle m,\langle c, z\rangle\rangle \in\left\{z_{0}, \ldots, z_{m}\right\} \subseteq \lambda x \cdot x x
$$

We produce such a fixed point for an arbitrary $c$, with $m=0$. That is, we consider a fixed point $z=\langle 0,\langle c, z\rangle\rangle$ (i.e. $z=\vartheta_{1} " 0 \cup \vartheta_{2} \vartheta_{1} " c \cup \vartheta_{2}^{2} " z$ ). Focusing on the existence of a $z$, which codes the pair $\langle c, z\rangle$ for an arbitrary set $c$, allows us to reduces the cases one would need to check (i.e. that each member of $\left\{z_{0}, \ldots, z_{m}\right\}$ is also a member of $\lambda x . x x$ ). In this case, as $z$ is the unique member of $\{z\}$ and $z=\langle c, z\rangle$, the condition that $z \in \lambda x . x x$ is satisfied automatically. So it remains to prove the existence of such a pair, for an arbitrary $c$.

The Quine pairing function (or, in this case, the Quine tripling function) is continuous in each variable, when NF is viewed as a cpo [10]. Thus, it has a least fixed point for an arbitrary c. An example of a fixed point, though not necessarily the least, is:

$$
z=\{\emptyset,\{0,1\},\{0,1,2,3\}, \ldots\} \cup \bigcup_{x \in c}\left\{\vartheta_{2} \vartheta_{1}^{‘} x, \vartheta_{2}^{3} \vartheta_{1} x, \vartheta_{2}^{5} \vartheta_{1} ‘ x, \ldots\right\}
$$

Remark (Evaluation). We construct one final combinator, as promised at the end of Section 5.2, which corresponds to cartesian closure of $\mathcal{R}$, the category of continuous retracts.

$$
\begin{align*}
\text { eval } & =\lambda u . \pi_{1}(u)\left(\pi_{2}(u)\right)  \tag{eval}\\
& =\left\{\langle n,\langle c, x\rangle\rangle \mid c \in f \operatorname{st}\left(\left\{x_{1}, \ldots, x_{n}\right\}\right)\left(\operatorname{snd}\left(\left\{x_{1}, \ldots, x_{n}\right\}\right)\right)\right\} \\
& =S(f s t)(\text { snd })
\end{align*}
$$

### 5.4 NF With Surjective Sequences

The implementation of finite sequences as nested Quine pairs can be extended to streams by recursion. Following Forster's implementation streams-as-sets (see [10]), we define a (canonical) homogeneous projection function $\pi: V^{\omega} \rightarrow V$, yielding an interpretation of sets-as-streams. In doing so, however, we uncover some complications. The standard implementation of sequences implies:

$$
\forall n \in N . V^{n}=V
$$

But, extended to streams, this implementation contains a redundancy:

$$
V^{\omega} \neq V
$$

This motivates the introduction of Quine sequences (Definition 5.57), producing a unique correspondence between sets and streams: $V^{\omega}=V$. By then forming the "quotient" $V^{\omega} / \sim_{\omega}$, where $\sim_{\omega}$ is the identity ${ }^{38}$

$$
\left\langle x_{0}, \ldots, x_{n}, \emptyset, \emptyset, \emptyset, \ldots\right\rangle \sim_{\omega}\left\langle x_{0}, \ldots, x_{n}\right\rangle
$$

we obtain an extended version of Scott's second conjecture $\left(V^{*}=V\right)$ :

$$
V=V^{* *} \equiv V^{*} \cup V^{\omega}
$$

While this is pleasing, at some levels, the extension of multi-relations to include $\omega$ sequences has nontrivial implications for the corresponding combinatory algebra of $\omega$ continuous functions.

### 5.4.1 Quine Pairs, A Further Look

To work "practically" within a given model of set theory requires the existence of ordered pairs. Yet, relatively little has been written about their various methods of implementation [59]. Wiener-Kuratowski are an intuitive and effective choice for $\mathrm{ZF}(\mathrm{C})$ but, due

[^150]to inhomogeneity, are ill-suited to stratified theories. Implementation of Quine pairs requires the existence of a partition of the universe $V=A \cup B$, along with a pair of bijective functions $V \rightarrow A$ and $V \rightarrow B[10]{ }^{39}$ This abstract property is implied by a more familiar one: the existence of a Dedekind infinite set (or natural numbers object) ${ }^{40}$ Once this condition is satisfied, however, Quine pairs provide surjective (type-level) pairing for any set theory in which they can be implemented.

## Formation of Sequences

$$
\text { Quine Pairs } \Longrightarrow V \times V=V
$$

More generally:

$$
V \times V=V \Longrightarrow \forall n \cdot V^{n}=V
$$

Sequences are implemented by iterative nesting of pairs (e.g. a 3 -tuple $\langle x, y, z\rangle$ is formed as $\langle z,\langle y, z\rangle\rangle)$. Accordingly, given any set $x$ and $n \in N$, there is a straightforward interpretation of $x$ as an $n$-sequence.

Example 5.47 (3-tuples). The 3 -tuple $\langle x, y, z\rangle$ is defined using the familiar $\vartheta_{0}$ and $\vartheta_{1}$ functions.

$$
\langle x,\langle y, z\rangle\rangle=\vartheta_{0} " x \cup \vartheta_{1} "\left(\vartheta_{0} " y \cup \vartheta_{1} " z\right)
$$

A similar method defines the three projection functions:

$$
\begin{aligned}
\pi_{1}^{3}(x) & =\left\{y \mid \vartheta_{0}(y) \in x\right\} \\
& =\{y \mid \exists z \in x .0 \notin z \wedge z \backslash N=y \backslash N \wedge \forall n \in N . n \in y \Longleftrightarrow n+1 \in z\} \\
\pi_{2}^{3}(x) & =\left\{y \mid \vartheta_{1} \vartheta_{0}(y) \in x\right\} \\
& =\{y \mid \exists z \in x .0 \in z \wedge 1 \notin z \wedge z \backslash N=y \backslash N \wedge \forall n \in N . n \in y \Longleftrightarrow n+2 \in z\} \\
\pi_{3}^{3}(x) & =\left\{y \mid \vartheta_{1} \vartheta_{1}(y) \in x\right\} \\
& =\{y \mid \exists z \in x .\{0,1\} \subseteq z \wedge z \backslash N=y \backslash N \wedge \forall n \in N . n \in y \Longleftrightarrow n+2 \in z\}
\end{aligned}
$$

[^151]Lemma 5.48. For any $y \in x$, where $x$ is interpreted as an $n$-tuple $\left\langle x_{0}, \ldots, x_{n-1}\right\rangle, y$ corresponds to an element of a unique member, $x_{j}$, of the sequence. We define two functions int and ext ${ }_{n}$ :

$$
\begin{aligned}
& \operatorname{int}(y)=\text { Greatest } m \in N \text { such that }\{l \mid l<m\} \subseteq y \\
& \operatorname{ext}_{n}(y)=\operatorname{int}(y \cap\{0, \ldots, n-2\})
\end{aligned}
$$

Therefore:

$$
\operatorname{ext}_{n}(y)=j \Longleftrightarrow y \in x_{j}
$$

As a corollary, projection functions are not independent of the chosen length of the sequence corresponding to $x$.

Corollary 5.49. For any set $x$, consider the corresponding $n$ and $n+1$-length sequences: $\left\langle x_{0}, \ldots, x_{n-1}\right\rangle$ and $\left\langle x_{0}, \ldots, x_{n}\right\rangle . \pi_{i}^{n-1}$ and $\pi_{i}^{n}$ denote the "ith" projection functions of the $n$ and $n+1$ length sequences, respectively. We obtain the result:

$$
\forall n . \pi_{n-1}^{n-1} \neq \pi_{n-1}^{n}
$$

This extends to $\pi_{n-1}^{n-1} \neq \pi_{n-1}^{m}$ for any length $m \geq n$.

Proof. By Lemma 5.48, for all $i<n, y \in \pi_{i}^{n}(x)$ if and only if $\{0,1, \ldots i-1\}$ is the longest initial segment of $N$ that is also a subset of $y$. However, in the case, $i=n\{0,1, \ldots i-1\}$ may be a proper subset of the longest initial segment.

The latter result is not a weakness per se, but it does imply:

$$
x_{n-1}^{n-1} \neq x_{n-1}^{n} \cup x_{n}^{n}
$$

Extending sequences is not equivalent to partitioning the final member of the previous sequence $\stackrel{41}{\square}$

[^152]Remark (Generalizing Sequences to Successor Ordinals). Given any concrete $n \in N$, iteration of the Quine pairing operations forms a homogeneous projection function (in its un-curried form):

$$
\pi^{n}:\{0, \ldots, n\} \times V \rightarrow V
$$

We can generalize sequence formation to any successor ordinal $\alpha \in N O$.
Lemma 5.50. Given some successor ordinal $\alpha$, there exists a homogeneous projection function:

$$
\pi^{\alpha}:\{\beta \mid \beta \leq \alpha\} \times V \rightarrow V
$$

associating a unique $\alpha$-sequence with each set. Informally, we obtain an identity:

$$
V^{\alpha}=V
$$

Proof. We first consider a generalisation of Quine's $\vartheta$-functions. Given any set $x$, we define two cases:

$$
\begin{align*}
& \vartheta_{\alpha}(\beta, x)=x \backslash N O \cup\{\gamma+1+\beta \mid \gamma \in x \cap N O\} \cup\{\delta \mid \delta<\beta\} \\
& \vartheta_{\alpha}(\alpha, x)=x \backslash N O \cup\{\gamma+\alpha \mid \gamma \in x \cap N O\} \cup\{\delta \mid \delta<\alpha\}
\end{align*}
$$

It remains to show that each set $y$ determines a unique $\beta \in N O$ and $x$, such that:

$$
\vartheta_{\alpha}(\beta, x)=y
$$

Define $\operatorname{int}(y)$ as the largest $\beta \in N O$ such that the initial sequence of $N O$ determined by $\beta$ is a subset of $y$ :

$$
\operatorname{int}(y)=\beta \Longleftrightarrow \prec_{\beta}=\cup\left\{\prec_{\gamma} \mid \forall \delta<\gamma \cdot \delta \in y\right\}
$$

We can then define $\beta$ as:

$$
\beta \equiv \max \{\operatorname{int}(y), \alpha\}
$$

Immediately, we obtain a unique set $x$, such that $\vartheta_{\alpha}(\beta, x)=y$. Again, we must consider two cases:

$$
\begin{align*}
& x=y \backslash N O \cup\{\gamma-(\beta+1) \mid \gamma \in x \cap\{\delta \mid \delta>\beta\}\} \\
& x=y \backslash N O \cup\{\gamma-\alpha \mid \gamma \in x \cap\{\delta \mid \delta \geq \alpha\}\}
\end{align*}
$$

It is clear that the two operations we have defined:

$$
\{\beta \mid \beta \leq \alpha\} \times V \rightarrow V \text { and } V \rightarrow\{\beta \mid \beta \leq \alpha\} \times V
$$

are mutually inverse.

As in Corollary 5.49, it is not generally the case that: $\pi_{\alpha}^{\alpha}=\pi_{\alpha}^{\alpha+1}$.

We now consider the most natural, concrete example of sequence formation for a limit ordinal: the formation of streams (i.e. $\omega$-sequences) in NF.

## Streams and Maximal Sequences

We can formally associate sets-to-streams and streams-to-sets, by recursion.
Lemma 5.51 (Streams to Sets [10]). Given some $\omega$-sequence, represented as a function $f: N \rightarrow V$, we can define a set:

$$
\vec{f}=\bigcup\left\{\left(\left(\vartheta_{1}\right)^{n} \circ \vartheta_{0}\right) " f(n) \mid n \in N\right\}
$$

Notice $\vec{f}$ is one type below $f$ in any valid stratification.

Likewise, any set $x$ has a corresponding stream $\left\langle x_{0}, x_{1}, \ldots\right\rangle$. We prove this by constructing a projection function $\pi: N \times V \rightarrow V$, following the dual construction in [10].

Lemma 5.52 (Sets to Streams). Given a set $x$, we define the corresponding stream, $\left\langle x_{0}, x_{1}, \ldots\right\rangle$, by constructing a function $f: N \rightarrow V$, where $f(n)=x_{n}$.

Proof. The original projection functions for Quine pairs are denoted:

$$
\begin{aligned}
& \bar{\pi}_{0}: x \mapsto\left\{y \mid \vartheta_{0}(y) \in x\right\} \\
& \bar{\pi}_{1}: x \mapsto\left\{y \mid\left(\vartheta_{0}(y) \cup\{0\}\right) \in x\right\}
\end{aligned}
$$

We define the function $g$ by recursion, where $g(0)=x$.

$$
g=\bigcap\left\{y \mid\langle 0, x\rangle \in y \wedge \forall\langle n, z\rangle \in y \cdot\left\langle n+1, \bar{\pi}_{1}(z)\right\rangle \in y\right\}
$$

From this we can define a canonical projection function for $\left\langle x_{0}, x_{1}, \ldots\right\rangle$, where $\pi_{x}(n)=x_{n}$ :

$$
\pi_{x}: N \rightarrow V ; \pi(n)=\bar{\pi}_{0} \circ g(n)
$$

Notice $\pi_{x}$ is one type above $x$ in any valid stratification.

Definition 5.53 (Canonical Projection). Taking $\pi_{x}$, as defined above, and ranging over $V$, we obtain a homogeneous function:

$$
\pi: N \times V \rightarrow V
$$

Defined by the action $\langle n, x\rangle \mapsto \pi_{x}(n)$.

## Complications with Nested Quine Pairs

Unfortunately, the standard method for coding finite sequences (nested Quine pairs) encounters complications, when extended to streams. Despite the fact that $V^{n}=V$ for any finite $n{ }^{[22}$ we do not obtain an exact equivalence: $V^{\omega}=V$.

Lemma 5.54. Let $\operatorname{seq}(x)$ denote the stream corresponding to $x$ :

$$
\forall x, y \cdot N \subseteq y \Longrightarrow \operatorname{seq}(x)=\operatorname{seq}(x \cup\{y\})
$$

Proof. The "coding" sets are the initial sequences of $N$. This includes, of course, $N$ itself. But each entry $x_{n}$ of $\operatorname{seq}(x)$ corresponds to a finite (' $n-1$ '-length) initial sequence of $N$ - more generally, to a successor ordinal. If the longest initial sequence of $N$ contained in an element $y \in x$ is a limit ordinal (i.e. $N \subseteq y$ ), then it does not correspond to a member of any "entry" in $\operatorname{seq}(x)$.

From Lemma 5.54, we obtain an immediate corollary that no stream is coded by a unique set.

[^153]Corollary 5.55. Given a stream $\vec{x}$, consisting of sets in a model of NF, coded by a set $x$, there exist sets distinct from $x$, also coding $\vec{x}$.

Example 5.56. $\operatorname{seq}(\{N\})=\operatorname{seq}(\emptyset)=\langle \rangle$.

One solution for this might be to move every entry $x_{n}$ of $\operatorname{seq}(x)$ to $x_{n+1}$ (we can easily adjust the formal coding to do this) and assign to $x_{0}$ the set whose members are coded by $\{y \mid y \in x \wedge N \subseteq y\}$. But this does not allow us to code any stream whose first entry has a non-empty intersection with $N$. In addition, we need pairwise disjointness between the collections of sets coding the universe of sets, at each entry of the sequence (i.e. the respective collections of sets, required for coding each set at the entries corresponding to $\pi(i,-)$ and $\pi(j,-)$, where $i \neq j$, must be disjoint).

### 5.4.2 Quine Sequences

We propose the following definition, from which we recover an exact correspondence between sets and streams - i.e. a surjective $\omega$-sequencing function that, as in the finite case, is injective.

Definition 5.57. We define the Quine sequence functions:

$$
\begin{aligned}
& \vartheta_{0}{ }^{\prime} x=\{2 n+1 \mid n \in x \cap N\} \cup\{z \mid z \in x \backslash N\} \cup 2 N \\
& \vartheta_{1} ‘ x=\{n+1 \mid n \in x \cap N\} \cup\{z \mid z \in x \backslash N\} \\
& \vartheta_{2}{ }^{‘} x=\{2 n+1 \mid n=0 \wedge n \in x\} \cup\{n+2 \mid n>0 \wedge n \in x\} \\
& \cup\{z \mid z \in x \backslash N\} \cup\{0\} \\
& \vartheta_{3}{ }^{‘} x=\{2 n+1 \mid n \leq 1 \wedge n \in x\} \cup\{n+3 \mid n>1 \wedge n \in x\} \\
& \cup\{z \mid z \in x \backslash N\} \cup\{0,2\} \\
& \cdots \\
& \vartheta_{m}{ }^{\prime} x=\{2 n+1 \mid n \leq m-2 \wedge n \in x\} \cup\{n+m \mid n>m-2 \wedge n \in x\} \\
& \cup\{z \mid z \in x \backslash N\} \cup\{0,2, \ldots, 2(m-2)\}
\end{aligned}
$$

For a stream of sets $x_{0}, x_{1}, \ldots$, viewed as a mapping $f: N \rightarrow V$, we form the corresponding set:

$$
\vartheta_{0} " f(0) \cup \vartheta_{1} " f(1) \cup \ldots \cup \vartheta_{n-1} " f(n-1) \ldots
$$

An equivalent, more formal definition of Quine sequences can be given by recursion.
Definition 5.58 (Sequences by Recursion). First, we define an operation $\downarrow$ and a (homogeneous) function $\tau{ }^{43}$

$$
\begin{align*}
& \downarrow x=n \Longleftrightarrow n \in 2 N \backslash x \wedge \forall m \in 2 N . m<n \Longrightarrow m \in x \\
& \tau^{\prime} x=\{n \mid n \in 2 N \wedge m \leq \downarrow x\} \cup\{z \mid z \in x \backslash N\} \\
& \cup\{n \mid n \in x \wedge n<\downarrow x+2 \wedge \exists m \cdot n=2 m+1\} \\
& \cup\{n+1 \mid n \geq \downarrow x+2\}
\end{align*}
$$

We can give a recursive definition of the Quine sequence functions:

$$
\begin{gathered}
\vartheta_{0} ‘ x=\{2 n+1 \mid n \in x \cap N\} \cup\{z \mid z \in x \backslash N\} \cup 2 N \\
\vartheta_{1}{ }^{‘} x=\{n+1 \mid n \in x \cap N\} \cup\{z \mid z \in x \backslash N\} \\
\text { for } n \geq 2 \text { we define: } \\
\vartheta_{n} ‘ x=\tau^{n-1 ‘} \vartheta_{1} ‘ x
\end{gathered}
$$

For a stream of sets $\left\langle x_{0}, x_{1}, \ldots\right\rangle$, viewed as a map $f: N \rightarrow V$, we form the corresponding set:

$$
\vartheta_{0} " f(0) \cup \vartheta_{1} " f(1) \cup \bigcup\left\{\tau^{n-1} \vartheta_{1} " f(n) \mid n \geq 2\right\}
$$

Similar to those we introduced for streams, in Lemma 5.52, Quine sequences have canonical projection functions. Using Quine sequences, we obtain an identity between sets and $\omega$-sequences.

Lemma 5.59. Quine Sequences imply the identity: $V^{\omega}=V$.

[^154]Remark (Impact of Sequence Formation on Multi-Relation Models). In a multi-relation model of $\lambda$-calculus, the utility of sequences is the coding of finite sets. In this context, there is no obvious "advantage" of $V^{\omega}=V$. An infinite sequence can code a finite set, and any finite set has a multitude of sequences, serving as codes. Thus, while the impact of an injective coding operation on a model of $\lambda$-calculus maybe interesting, it is orthogonal to whether or not our (surjective) $\omega$-sequence function is injective ${ }^{44}$

### 5.4.3 A "Cumulative Hierarchy" of Sequences

Lemma 5.50 showed that, given any successor ordinal $\alpha$, we could form mutually inverse sequence and projection operations. From this we obtained an identity:

$$
V^{\alpha}=V
$$

In particular, for any $n \in N, V^{n}=V$. The introduction of Quine sequences extended this identity to include $\omega$-sequences. What interests us, however, is something more like a "cumulative hierarchy" of sequences:

$$
V^{*}=\bigcup_{n \in N} V^{n}
$$

In this section, we seek an exact equivalence between sets and the cumulative collection of countable sequences:

$$
V=V^{* *} \equiv V^{\omega} \cup \bigcup_{n \in N} V^{n}
$$

## Streams as (Possibly) Finite Sequences

Much like $\lambda$-abstraction, the operations associating sets and streams are external (i.e. inhomogeneous) but, nevertheless, mutually inverse. Therefore, we obtain a homogeneous correspondence:

$$
V^{\omega} \rightarrow V
$$

[^155]Just as we extend a function $f: X \rightarrow Y$ to a partial function $f: V \rightarrow V$, which maps sets in $V \backslash X$ to $\emptyset$, we consider an extension of finite sequences to streams. We can make the identification $\sqrt[45]{45}$

$$
\left\langle x_{0}, \ldots, x_{n}\right\rangle \sim_{\omega}\left\langle x_{0}, \ldots, x_{n}, \emptyset, \emptyset, \ldots\right\rangle
$$

Using $\sim_{\omega}$, we obtain a canonical association of each set in NF with a canonical (possibly infinite) sequence. The length of this sequence corresponds to a function:

$$
\text { extent : } V \rightarrow N \cup\{\omega\}
$$

The definition given in Lemma 5.48 can be extended to Quine sequences, where the extent (i.e. length) of the sequence corresponding to a given set, is given by the longest, proper initial segment of 2 N contained in one of its members.

Definition 5.60. $\operatorname{extent}(x)=j+2$ precisely when the initial segment $\{0, \ldots, 2 j-2\}$ is the longest proper initial segment of $2 N$, contained in the set:

$$
\{y \cap 2 N \mid y \in x\}
$$

We need to include some special cases:

$$
\begin{array}{ll}
\forall y \in x .2 N \subseteq y & (\operatorname{extent}(x)=1) \\
\exists y \in x . y \cap N=\emptyset \wedge \forall y \in x . y \cap N=\emptyset \vee 2 N \subseteq y & (\operatorname{extent}(x)=2) \\
\forall m \in 2 N . \exists y^{\prime} \in[\{y \cap 2 N \mid y \in x\} \backslash 2 N] . m \subsetneq y^{\prime} & (\operatorname{extent}(x)=\omega)
\end{array}
$$

The unique set $x$ with $\operatorname{extent}(x)=0$ is $\emptyset$.

Remark (Uniqueness of Projections). Notice that we not only have a canonical sequence associated to each set, but we now have a uniform collection of projection functions ${ }^{46}$

$$
\pi_{i}^{\omega}: V^{\omega} / \sim_{\omega} \rightarrow V
$$

[^156]independent of the length, $\operatorname{extent}(x)$, of the sequence corresponding to each set $x{ }^{\boxed{47}}$

Definition 5.61. For a set theory with universe $V$, we define $V^{* *}$ as the collection of countable sequences of sets. $P^{`} V^{* *}$ is the set of countable multi-relations.

Lemma 5.62. Quine Sequences imply the identity: $V^{\omega} / \sim_{\omega}=V^{* *}$.

Together, Lemma 5.59 and 5.62 determine an exact equivalence:

$$
V=V^{* *}
$$

### 5.4.4 Application of Streams

Using $\pi: N \times V \rightarrow V$ (see Definition 5.53), we can define a homogeneous application function.

Definition 5.63. The operation $a p p_{\omega}: V \times V \rightarrow V$ denotes application for NF with streams (implemented with Quine Sequences):

$$
\operatorname{app}_{\omega}(x, y)=\{z \mid \exists a \in x \cdot \pi(0, a)=z \wedge\{b \mid \exists n \in N \backslash\{0\} \cdot \pi(n, a)=b\} \subseteq y\}
$$

In order to simplify future notation, we define the (inhomogeneous) operation $\chi 4^{48}$

$$
\chi(a)=\{b \mid \exists n \in N \backslash\{0\} \cdot \pi(n, a)=b\}
$$

Thus we can (notationally) simplify our definition of $a p p_{\omega}$ :

$$
\operatorname{app}_{\omega}(x, y)=\{z \mid \exists a \in x \cdot \pi(0, a)=z \wedge \chi(a) \subseteq y\}
$$

[^157]As $\lambda$-abstractions are "look-up tables," we do not have to worry about a canonical representation of a given set, as a sequence 4 It is useful to form an operation mapping a finite set $x$ to its set of "codes."

Definition 5.64. The operation, which takes any set $x$ and maps it to the collection of sets $a$ such that $\{z \mid \exists n \in N . \pi(n, a)=z\}=x$, is denoted $\gamma(-)$

$$
\gamma(x)=\{a \mid \forall n \in N . \pi(n, a) \in x . \wedge . \forall y \in x . \exists n \in N . \pi(n, a)=y\}
$$

This operation is stratified and, perhaps surprisingly, homogeneous 50

A couple observations, regarding $a p p_{\omega}$ and $\chi$ :

1. $\chi^{"} V=(F i n \cup|\omega|) \backslash\{\emptyset\}$.
2. $\operatorname{app}_{\omega}(\emptyset, x)=\emptyset=\operatorname{app}_{\omega}(x, \emptyset)$

Using our analogy with multi-relations, the first observation $\left(\emptyset \notin \chi^{\prime} V\right)$ implies that, while we can form potentially infinite sequences, we cannot define one element sequences $\langle z\rangle$. As such, we are unable to form elements, which will be "constants" under application 5 This relates to the second property, which states that $\emptyset$ is a two-sided fixed point of $a p p_{\omega}$.

## $\lambda$-Abstraction.

Question. What is the appropriate form of $\lambda$-abstraction? To what extent does this imply $\beta$-equivalence between $\omega$-continuous functions and their $\lambda$-abstracts?

[^158]It will be helpful to distinguish between functions defined by their action on finite sets and the larger class of functions defined by their action on (possibly infinite) sets with cardinality less than or equal to $\omega$. We refer to these functions, respectively, as continuous and $\omega$-continuous $\sqrt{52}$

$$
\begin{array}{lr}
\forall x . f(x)=\bigcup\left\{f(s) \mid s \in P_{\aleph_{0}}(x)\right\} & (f \text { is continuous }) \\
\forall x . g(x)=\bigcup\{g(t)|t \subseteq x \wedge| t \mid \preceq \omega\} & (g \text { is } \omega \text {-continuous })
\end{array}
$$

We consider two candidates for the $\lambda$-operator, in the presence of $a p p_{\omega}$. The first respects Scott's (now standard) definition of continuous as "finite character" 57]. The second considers our expansion to countable multi-relations, $V^{* *}$, which permits coding of (possibly) infinite sets.

Definition 5.65 (Possible $\lambda$-operators).

1. $\lambda x \cdot f=\{z \mid \chi(z) \in$ Fin $\wedge \pi(0, z) \in f \circ \chi(z)\}$
2. $\lambda_{\omega} x . f=\{z \mid \pi(0, z) \in f \circ \chi(z)\}$

The latter is a more appropriate definition of $\lambda$-abstraction, with $\omega$-sequences. The following lemma summarizes why.

## Lemma 5.66.

1. For any function $f: V \rightarrow V, \lambda x . f \subseteq \lambda_{\omega} x . f$.
2. Any continuous function $f$ is $\beta$-equivalent to $\lambda_{\omega} x$. $f$.
3. $\beta$-equivalence fails between an $\omega$-continuous function $g: V \rightarrow V$ and $\lambda x . g$.

Furthermore, our application operation $a p p_{\omega}$ is itself $\omega$-continuous, and would not be $\beta$-equivalent to $\lambda x y \cdot \operatorname{app}_{\omega}(x, y)$.

[^159]While $\lambda x y$. app $_{\omega}$ is not $\beta$-equivalent to app $_{\omega}$, it is self-evident from Definition 5.63 that it exists as an object (i.e. combinator) in our model. We referred to this situation in Section 5.1 as incoherence between abstraction and application. In Section 5.5 we will see that coherence is not the only concern. When application is defined as app $\omega_{\omega}$, $\lambda_{\omega} x y z . x(y)(x(z))$ becomes $\omega$-continuous. Proving the existence of an $S$-combinator, therefore, requires a (weak) form of choice.

## Combinatory Algebra of Streams

Just as the calculus of distributive functions, for NF with Quine pairs, can be thought of as the "natural" calculus of functions for NF, the calculus of $\omega$-continuous functions can be thought of as the natural calculus for NF with Quine sequences.

The $I$ and $K$ combinators do not invoke application, so there is no distinction between (finite) continuity and $\omega$-continuity. If we assume countable choice, we obtain $\omega$-continuity of $S$ (Proposition 5.95). But, even assuming countable choice, the "natural" combinatory algebra of NF with Quine Sequences is not a model of standard (untyped) $\lambda$-calculus.

## An Error Term

The set $\emptyset$ acts as an error term, which is a two-side fixed point of $a p p_{\omega}$.

$$
\forall x \cdot \operatorname{app}_{\omega}(x, \emptyset)=\operatorname{app}_{\omega}(\emptyset, x)=\emptyset
$$

In other words, an "error" state always evaluates to "break." The semantic interpretation is that an error state does not contain "data" to which a program can be applied. The machine will run, but even the program corresponding to $\bar{K}(x)$, the $\omega$-continuous fragment of $K$ applied to some data $x$, will return an error. Thus, in general, continuous functions are only $\beta$-equivalent to their abstracts modulo $\emptyset$. A natural question is then: does each $\lambda$-abstraction correspond to a unique function?

Given two continuous functions, $\tau$ and $\tau^{\prime}$, with:

$$
\lambda_{\omega} x \cdot \tau[x]=\lambda_{\omega} x \cdot \tau^{\prime}[x]
$$

both $\tau(\emptyset)$ and $\tau^{\prime}(\emptyset)$ must be contained in:

$$
\bigcap_{y \in V \backslash \emptyset} a p p p_{\omega}\left(\lambda_{\omega} x \cdot \tau[x], y\right)
$$

the set of "default returns" for $\tau$ and $\tau^{\prime}$, interpreted as programs. Thus, with each $\lambda$-abstraction, we can associate a maximal continuous function.

Lemma 5.67. Given a set $x$, the maximal completion of $x$ :

$$
1 x=\lambda_{\omega} y \cdot a p p_{\omega}(x, y)
$$

is equivalent to the $\lambda$-abstraction of a set of ( $\omega$-)continuous functions, equivalent on $V \backslash\{\emptyset\}$ and determined by the intersection:

$$
\bigcap_{y \in V \backslash \emptyset} \operatorname{app}_{\omega}(1 x, y)
$$

Lemma 5.67 implies that any $(\omega$-)continuous function $\tau$, with no default values, is $\beta$ equivalent to $\lambda_{\omega} x . \tau[x]$.

Corollary 5.68. For any ( $\omega$-)continuous function $\tau$ :

$$
\bigcap_{y \in V \backslash \emptyset} \tau(y)=\emptyset \Longrightarrow \tau \sim_{\beta} \lambda_{\omega} x . \tau[x]
$$

Remark (Handling $\emptyset$ ). There are certainly ways to handle the complications, regarding the coding of "default values" with streams. One could, for example, add a step at the end of the de-coding process, which maps each $n \in N$ to $n-1$ and maps 0 to $\emptyset$. At that point, one could interpret $\langle x, 0\rangle$ as $\langle x\rangle$, without giving up the ability to express any pairs of the form $\langle x, n\rangle$, where $n \in N$.

### 5.5 Generalizing Sequences to Codes

The models of pure $\lambda$-calculus developed in [55] and [57] are referred to as continuous models [53. The word "continuous" is appropriate in a formal, topological sense - for injective $T_{0}$-spaced (equally, continuous lattices), specifically.
[58] moves toward the study of function spaces for unstructured ${ }^{\sqrt{53}}$ collections of objects (i.e. models of set theory). ${ }^{54}$ While a more general, unstructured context does not preclude the existence of a relevant topology, here we emphasize the coding aspect of set theoretic models of combinatory algebra on $\mathcal{A}{ }^{55}$

$$
\alpha, \chi: \mathcal{A} \rightarrow \mathcal{A}
$$

with corresponding application and abstraction operations:

$$
\begin{array}{ll}
\operatorname{app}_{\chi}(x, y)=\{z \mid \exists w \in x . \alpha(w)=z \wedge \chi(w) \subseteq y\} & (\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}) \\
\lambda_{\chi} x \cdot f=\{z \mid \alpha(z) \in f \circ \chi(z)\} & (\mathcal{A} \Rightarrow \mathcal{A} \rightarrow \mathcal{A})
\end{array}
$$

Remark (Combinators Remain $\subseteq$-Continuous). For any set $x \in \mathcal{A}$, the definition of app $_{\chi}$ implies:

$$
\operatorname{app}_{\chi}(x,-): \mathcal{A} \rightarrow \mathcal{A}
$$

is a $\subseteq$-continuous function. Regardless of which sets are "coded" by $\chi$, no function can be $\beta$-equivalent to some combinator, which is not itself $\subseteq$-continuous. Thus, even assuming only the most elementary structure on $\mathcal{A}$, the combinatory algebra satisfies a continuity principle. But this does not necessarily mean $\chi$ is $\subseteq$-continuous, itself. We also do not assume coding is a form of "size" condition - a set $x$ may be of the form $\chi(z)$, but there could easily exist some $y$, such that $|y| \prec|x|$ and $y$ is not in the image of $\chi$.

The goal of this section is to develop a general framework, within which we can better understand the relationship between those sets that are "coded" (e.g. finite sets, in
${ }^{53}$ Of course, even a model of set theory has the structure of a partial order, induced by the $\subseteq$-relation.
${ }^{54}$ Plotkin investigated "graph models" of untyped $\lambda$-calculus shortly after Scott's initial publication of Continuous Lattices [44. His work (and Scott's work, independently, in [56) is the first general study of models of combinatory algebra in set theory, of which multi-relation models can be seen as a special case. The current section departs from Plotkin, for better or worse, in de-emphasizing the role of continuous lattices and, in some sense, placing greater emphasis on the role of comprehension. In this way, while Section 5.5 proves things in greater generality, it is always with an eye toward application in NF.
${ }^{55}$ We have already encountered an example of this in Section 5.3 . (See Definitions 5.34, 5.35 and 5.36 ) In general, we assume $\alpha$ is surjective, as in the case of Definition 5.34
the standard multi-relation model [58]) and those ( $\subseteq$-continuous) functions that are $\beta$ equivalent to their corresponding $\lambda$-abstracts. Primarily, we focus on how $\alpha$ and $\chi$ relate to the standard combinators: $S, K$ and $I$. To better understand the necessary coding conditions for the existence of a given combinator in our model, we introduce coding trees. Each set $x \in \mathcal{A}$ determines a canonical tree of "coded" sets, obtained by iteration of $\chi$ and $\alpha{ }^{56}$

Example 5.69 (Distributive Coding). In the distributive case (Section 5.2) the coding operations are just the first and second projections of the Quine pairing operation. For an arbitrary set $x$, we give the first two levels of the resulting coding tree:


Example 5.70 (Multi-Relation Coding). In the multi-relation model, we have a series of operations:

$$
\begin{aligned}
& \text { seq }: \mathcal{A} \rightarrow \mathcal{A}^{*}: x \mapsto \vec{x} \\
& \pi_{0}: \mathcal{A}^{*} \rightarrow \mathcal{A}: \vec{x}=\left\langle x_{0}, \ldots, x_{n}\right\rangle \mapsto x_{0} \\
& \pi_{1, n}: \mathcal{A}^{*} \rightarrow \mathcal{A}^{*}: \vec{x}=\left\langle x_{0}, \ldots, x_{n}\right\rangle \mapsto\left\langle x_{1}, \ldots, x_{n}\right\rangle \\
& \text { set }: \mathcal{A}^{*} \rightarrow \mathcal{A}: \vec{x} \mapsto\left\{z \mid \exists x_{i} \in \vec{x} . z=x_{i}\right\}
\end{aligned}
$$

From these operations, we obtain codes from the composites $\sqrt{57}$

$$
\begin{aligned}
& \alpha \equiv \pi_{0} \circ s e q: \mathcal{A} \rightarrow \mathcal{A} \\
& \chi \equiv \operatorname{set} \circ \pi_{1, n} \circ s e q: \mathcal{A} \rightarrow \mathcal{A}
\end{aligned}
$$

[^160]Each set $x$ in $\mathcal{A}$ is the root of a coding tree:


## Properties of Coding Trees

Similar to the $\epsilon$-tree representation of sets, one may wish to consider those systems where each coding tree is determined by a unique root. ${ }^{58]}$

Definition 5.71. We will refer to a coding (i.e. a pair of coding operations $\alpha$ and $\chi$, as defined above) as extensional if each set is uniquely determined by its coding tree. We can see, by induction, this is determined entirely by the initial branches of the coding tree. In other words:

$$
\forall x, y \cdot x=y \Longleftrightarrow(\alpha(x)=\alpha(y) \wedge \chi(x)=\chi(y))
$$

Example 5.72. The distributive coding is clearly extensional. The multi-relation coding is not.

We are also interested in what we will refer to (for lack of a better word) as the generative property. From an existing set $x$, we can form the first level of its corresponding tree by the operation:

$$
\langle\alpha, \chi\rangle: \mathcal{A} \rightarrow \mathcal{A} \times \mathcal{A}
$$

On the other hand, given a pair of sets $\langle y, z\rangle$, we can ask if it occurs as the first level of a coding tree, for some set $x{ }^{59}$ In terms of the coding operations, this is equivalent to the

[^161]fibre of $\langle\alpha, \chi\rangle$ over $\langle y, z\rangle$ being nonempty ${ }^{60}$ More generally, given the first $n$ levels of a coding tree, where each of its terminal nodes is decorated with a set, we say a tree has the generative property if it determines the existence of a root (i.e. a set in $\mathcal{A}$, whose canonical coding tree is equivalent to the one we have given).

Definition 5.73. Given an $n$-level coding tree, the terminal nodes are those (on the right) corresponding to $\chi, \chi \alpha, \ldots \chi \alpha^{n-1}$ and (the leftmost node) corresponding to $\alpha^{n}$.

Remark (Nested Sequences and Coding Trees). In a multi-relation model, the $\lambda$-abstract of an $n$-variable function consists of $n$-nested sequences. In the parlance of Example 5.70, $n$-nested sequences correspond to $n$-level coding trees.

Example 5.74 (Generative Coding Trees with Multi-relations). Definition 5.64 associated each set $x$ with the collection, $\gamma(x)$, of streams whose entries consist of (possibly multiple copies of) exactly the members of $x{ }^{61}$

$$
\gamma(x)=\{a \mid \forall n \in N . \pi(n, a) \in x . \wedge . \forall y \in x . \exists n \in N . \pi(n, a)=y\}
$$

The standard multi-relation model is generative. If we decorate the terminal nodes on the right hand side of some $n$-level coding tree (i.e those which would be of the form $\chi \alpha^{i}$ ) with sets $x$, such that $\gamma(x) \neq \emptyset$, the remaining (leftmost) terminal node (i.e. one of the form $\alpha^{n}$ ) can be decorated by any set and the result will imply the existence of a (finite) root, formed by concatenation of finite sequences.

Remark (Toward $\chi$-Continuity). The motivation for all this is to determine necessary and sufficient properties of our coding system, to $\lambda$-abstract specific functions. For example, if we wish to $\lambda$-abstract even the most basic function, $\lambda x . x$, we need $\alpha: \mathcal{A} \rightarrow \mathcal{A}$ to be surjective.

Now, instead of $\lambda x . x$, consider an arbitrary ( $\subseteq$-continuous) function $f: \mathcal{A} \rightarrow \mathcal{A}$ and a coding system $\langle\alpha, \chi\rangle$. Consider the minimal sub-collection $\mathcal{A}_{f} \subseteq \mathcal{A}$ of sets on which $f$ is determined:

$$
\forall x \in \mathcal{A} \cdot f(x)=\bigcup_{s \in \mathcal{A}_{f} \cap P(x)} f(s)
$$

[^162]$\mathcal{A}_{f} \subseteq \chi$ " $\mathcal{A}$ implies that $f$ is determined by its action on coded sets:
$$
\forall x \in \mathcal{A} \cdot f(x)=\bigcup_{s \in \chi \text { " } \mathcal{A} \cap P(x)} f(s)
$$

Such functions are said to be $\chi$-continuous.

The existence of a set (viewed as a combinator) in $\mathcal{A}$ that is $\beta$-equivalent to $f$ requires a stronger condition:

$$
\forall s \in \mathcal{A}_{f} \cdot \forall x \in f(s) \cdot \exists z \cdot \alpha(z)=f(x) \wedge \chi(z)=s
$$

Proving this corresponds to the existence of generative trees for each $s \in \mathcal{A}_{f} \cap \chi^{\prime} \mathcal{A}$ and each $x \in f(s)$, where the lefthand terminal node is $x$ and the righthand terminal node is $s{ }^{62}$

Remark (Concrete Coded Sets). Rather than focusing on specific functions, we may want to start by focusing on $\chi$ " $\mathcal{A}$, the entire collection of coded sets. In the generative case (e.g. multi-relations), any function determined by its action on coded sets is $\beta$ equivalent to some combinator in the multi-relation model (among such sets, the $\lambda$ operator identifies a canonical, maximal combinator). While such a condition does not hold, in general, it is of interest to consider the sub-collection of $\chi$ " $\mathcal{A}$ for which it does. In other words, we wish to determine those coded sets $z \in \chi^{\prime} \mathcal{A}$ where, for any function $f$ :

$$
f^{\prime} z=a p p_{\chi}\left(\lambda_{\chi} \cdot f, z\right)
$$

Definition 5.75. Any coded set $z$ (i.e $z \in \chi^{\prime \prime} \mathcal{A}$ ) satisfying the property:

$$
\forall y \exists w \cdot \alpha(w)=y \wedge \chi(w)=z
$$

is referred to as a concrete coded set.

Remark (Concrete Coded Sets Correspond to Generative Trees). For a coding tree with a single level (i.e. a pair of terminal nodes, from a common root), Definition 5.75 says: if the right-hand terminal node is decorated with a concrete coded set, the tree is generative. This property applies to any $n$-level coding tree: if the rightmost terminal nodes (i.e.

[^163]those corresponding to $\chi, \chi \alpha, \ldots \chi \alpha^{n-1}$ of the root) are decorated with concrete coded sets, the tree is generative, regardless of which set is chosen to decorate the leftmost terminal node (corresponding to $\alpha^{n}$ ) ${ }^{63}$

Any ( $\chi$-continuous) function determined by its action on concrete coded sets is $\beta$ equivalent to some set in $\mathcal{A} \cdot{ }^{64}$

### 5.5.1 Pre-Combinatory Algebra

Definition 5.76. A collection of sets $\mathcal{A}$ and a pair of coding operations $\alpha, \chi: \mathcal{A} \rightarrow \mathcal{A}$, where $\alpha$ is surjective, form what we refer to as a pre-combinatory algebra, classified by the following operations $\sqrt{65}$

$$
\begin{aligned}
& \gamma(x)=\{a \mid \chi(a)=x\} \\
& \operatorname{app}_{\chi}(x, y)=\{z \mid \exists w \in x . \alpha(w)=z \wedge \chi(w) \subseteq y\} \\
& \lambda_{\chi} x . f=\{z \mid \alpha(z) \in f \circ \chi(z)\}
\end{aligned}
$$

$\chi " \mathcal{A} \subseteq \mathcal{A}$ denotes the collection of all coded sets (i.e. sets of the form $\chi(y)$ ).
Definition 5.77. A pre-combinatory algebra $(\alpha, \chi, \mathcal{A})$, with the generative coding property, is referred to as a concrete pre-combinatory algebra.

Important: For the remainder of this section, any pre-combinatory algebra is assumed to be a concrete pre-combinatory algebra. As such, we may infer the existence of a set $w$, as the root of a coding tree whose terminal nodes are decorated with sets.

Definition 5.78. $P_{\chi}$ denotes the coded powerset:

$$
P_{\chi}^{‘} x=\{y \mid y \subseteq x \wedge \gamma(y) \neq \emptyset\}
$$

[^164]For NF, $P_{\chi}{ }^{`} V=\chi$ " $V$ and $\gamma^{\prime \prime} P_{\chi}{ }^{`} V$ gives a partition of the universe, induced by the fibre of $\chi$ over each coded set.

Definition 5.79. A function $f: V \rightarrow V$ is $\chi$-continuous if:

$$
\forall x . f^{\prime} x=\bigcup_{y \in P_{\chi}^{\prime} x} f^{\prime} y
$$

For any function $f$, we refer to $\bar{f}=\lambda_{\chi} x$.f as the $\chi$-continuous fragment of $f$.

By definition, $\operatorname{app}_{\chi}(-,-)$ is $\subseteq$-continuous. So we have the property:

$$
\operatorname{app}_{\chi}\left(\lambda_{\chi} x . f, y\right) \subseteq f^{\prime} y
$$

Therefore, $\chi$-continuity of $f$ is equivalent to:

$$
f^{\prime} y \subseteq \operatorname{app}_{\chi}\left(\lambda_{\chi} x . f, y\right)
$$

## $\chi$-Continuity of Standard Combinators

The primary motivation for the level of abstraction in a pre-combinatory algebra is to determine necessary and sufficient conditions for $\chi$-continuity of various combinators. As such, we have made minimal assumptions regarding $\alpha$ and $\chi$. Among the most basic properties we might ask is that $\chi$-application is $\chi$-continuous. In other words, we expect:

$$
A \equiv a p p_{\chi}(-,-): V \times V \rightarrow V
$$

to be $\beta$-equivalent to its $\chi$-continuous fragment:

$$
\begin{aligned}
\bar{A} \equiv & \lambda_{\chi} x y \cdot a p p_{\chi}(x, y): V \rightarrow(V \rightarrow V) \\
& =\left\{w \mid \exists s \in \chi(w) \cdot \alpha(s)=\alpha^{2}(w) \wedge \chi(s) \subseteq \chi \alpha(w)\right\}
\end{aligned}
$$

The equivalence is not directly implied by Definition 5.76. It requires the coding to be at least as strong as the distributive coding (Example 5.69).

Lemma 5.80 (A). $a^{p p} p_{\chi}$ is a $\chi$-continuous function if and only if the following condition holds:

$$
\begin{aligned}
\forall s \exists w \cdot \alpha^{2}(w) & =\alpha(s) \\
\wedge \chi(s) & =\chi \alpha(w) \\
\wedge \chi(w) & =\{s\}
\end{aligned}
$$

Proof. We need to prove:

$$
\forall x, y . \forall z . z \in \operatorname{app}_{\chi}(x, y) \Longleftrightarrow z \in \bar{A}(x)(y)
$$

Rephrase $\operatorname{app}_{\chi}(x, y)$ as:

$$
\operatorname{app}_{\chi}(x, y)=\alpha "\{w \in x \mid \chi(w) \subseteq y\}
$$

Thus, proving $\chi$-continuity of $A$ reduces to proving:

$$
\begin{aligned}
\forall x, y . \forall s \in x \cdot \chi(s) \subseteq y \Longleftrightarrow \exists w \cdot s & \in \chi(w) \\
& \wedge \chi(s) \subseteq \chi \alpha(w) \subseteq y \\
& \wedge \alpha^{2}(w)=\alpha(s) \wedge \chi(w) \subseteq x
\end{aligned}
$$

Given any set $s$, the existence of a set $w$ inducing the following coding tree is clearly sufficient to satisfy the above condition.


To see the existence of such a set is necessary, it suffices to consider: $\operatorname{app}_{\chi}(\{s\}, \chi(s))$.

Definition 5.81. The $\chi$-continuous fragments of $S, K$ and $I$ are denoted: $\bar{S}, \bar{K}$ and $\bar{I}$.

$$
\begin{align*}
& \lambda_{\chi} x \cdot x=\{w \mid \alpha(w) \in \chi(w)\}  \tag{I}\\
& \lambda_{\chi} x y \cdot x=\left\{w \mid \alpha^{2}(w) \in \chi(w)\right\}  \tag{K}\\
& \lambda_{\chi} x y z \cdot x(z)(y(z))  \tag{S}\\
& =\left\{w \mid \exists s \in \chi(w) \cdot \chi(s) \subseteq \chi \alpha^{2}(w)\right. \\
& \left.\wedge \forall t \in \chi \alpha(s) . \exists y \in \chi \alpha(w) . \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w)\right\}
\end{align*}
$$

Lemma 5.82 (I). A pre-combinatory algebra $(\alpha, \chi, \mathcal{A})$ has the I-combinator if and only $i f \underbrace{66}$

$$
\forall x \cdot \exists z \cdot \alpha(z)=x \wedge \chi(z)=\{x\}
$$

Lemma $5.83(\mathrm{~K}) .(\alpha, \chi, \mathcal{A})$ has the $K$-combinator if and only if the following condition holds:

$$
\begin{aligned}
\forall s \cdot \exists w \cdot \alpha^{2}(w) & =s \\
\wedge \chi(w) & =\{s\} \\
\wedge \chi \alpha(w) & =\emptyset
\end{aligned}
$$

Proof. We need to prove, given any pair of sets $x$ and $y$ :

$$
x \subseteq \bar{K}(x)(y)=a p p_{\chi}\left(a p p_{\chi}(\bar{K}, x), y\right)
$$

Recall $\bar{K}=\left\{w \mid \alpha^{2}(w) \in \chi(w)\right\}$, therefore $\bar{K}(x)(y)$ defines the following set:

$$
\bar{K}(x)(y)=\left\{s \mid \exists w \cdot \alpha^{2}(w) \in \chi(w) \subseteq x \wedge \alpha^{2}(w)=s \wedge \chi \alpha(w) \subseteq y\right\}
$$

To prove

$$
\forall x, y . s \in x \Longrightarrow s \in \bar{K}(x)(y)
$$

it clearly suffices that the following condition holds:

$$
\forall s \cdot \exists w \cdot \alpha^{2}(w)=s \wedge \chi(w)=\{s\} \wedge \chi \alpha(w)=\emptyset
$$

[^165]On the other hand, we can see this is a necessary condition, by considering $\bar{K}(\{s\})(\emptyset)$, for any set $s$. $\bar{K}(\{s\})(\emptyset)=\{s\}$ is equivalent to the condition:

$$
\exists w \cdot \alpha^{2}(w)=s \wedge \chi(w)=\{s\} \wedge \chi \alpha(w)=\emptyset
$$

## S-Combinator

It is less straightforward to determine the necessary and sufficient conditions for:

$$
\bar{S}(a)(b)(u)=a(u)(b(u))
$$

$\bar{S}$ applied to $a, b$ and $u$, results in the set:

$$
\begin{aligned}
& \bar{S}(a)(b)(u)= \\
&=\{x \mid \exists c \in \bar{S}(a)(b) \cdot \alpha(c)=x \wedge \chi(c) \subseteq u\} \\
&=\{x \mid \exists w \cdot \exists s \in \chi \chi(w) \cdot \chi(s) \subseteq \chi \alpha^{2}(w) \\
& \wedge \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w) \\
&\left.\wedge \alpha^{3}(w)=x \wedge \chi \alpha^{2}(w) \subseteq u \wedge \chi \alpha(w) \subseteq b \wedge \chi(w) \subseteq a\right\}
\end{aligned}
$$

$a(u)(b(u))$ results in the set:

$$
\begin{aligned}
\operatorname{app}_{\chi}\left(a p p_{\chi}(a, u), \operatorname{app} p_{\chi}(b, u)\right)=\{x \mid \exists s & \in a \cdot \alpha^{2}(s)=x \wedge \chi(s) \subseteq u \\
& \wedge \forall t \in \chi \alpha(s) . \exists y \in b \cdot \alpha(y)=t \wedge \chi(y) \subseteq u\}
\end{aligned}
$$

As the graph of a $\chi$-continuous fragment is contained in the graph of the original function:

$$
\bar{S}(a)(b)(u) \subseteq a(u)(b(u))
$$

Thus, proving $\chi$-continuity of $S$ is equivalent to proving the converse:

$$
\forall x . x \in a(u)(b(u)) \Longrightarrow x \in \bar{S}(a)(b)(u)
$$

$x \in a(u)(b(u))$ is equivalent to the conditions:

$$
\begin{align*}
\exists s \in & a \cdot \chi(s) \subseteq u  \tag{1a}\\
& \wedge \forall t \in \chi \alpha(s) . \exists y \in b \cdot \alpha(y)=t \wedge \chi(y) \subseteq u  \tag{1b}\\
& \wedge \alpha^{2}(s)=x \tag{1c}
\end{align*}
$$

For $x$ to be a member of $\bar{S}(a)(b)(u)$, conditions $(1 a)-(1 c)$ must imply:

$$
\begin{align*}
\exists w . s & \in \chi(w) \wedge \chi(w) \subseteq a  \tag{2a}\\
& \wedge \chi(s) \subseteq \chi \alpha^{2}(w) \subseteq u  \tag{2b}\\
& \wedge \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w)  \tag{2c}\\
& \wedge \chi \alpha(w) \subseteq b  \tag{2d}\\
& \wedge \alpha^{3}(w)=\alpha^{2}(s)=x \tag{2e}
\end{align*}
$$

It will also be helpful to consider alternative (equivalent) versions of conditions (1b) and (2c):

$$
\begin{align*}
& \chi \alpha(s) \subseteq \alpha "\{y \in b \mid \chi(y) \subseteq u\}  \tag{*}\\
& \chi \alpha(s) \subseteq \alpha "\left\{y \in \chi \alpha(w) \mid \chi(y) \subseteq \chi \alpha^{2}(w)\right\} \tag{*}
\end{align*}
$$

Any coding method strong enough to support the implication:

$$
\text { Conditions: } 1 a-1 c \Longrightarrow \text { Conditions: } 2 a-2 e
$$

will be sufficient for $\chi$-continuity of the $S$-combinator. Necessary and sufficient conditions for the above implication correspond to minimal assumptions regarding the coding system.

## A Reduced Form of $S$

To determine "minimal" coding properties, necessary for $\chi$-continuity of $a(u)(b(u))$, a good starting point is to consider the $\subseteq$-minimal subsets $a^{\prime} \subseteq a, b^{\prime} \subseteq b$ and $u^{\prime} \subseteq u$ such that:

$$
\forall a, b, u \cdot a(u)(b(u))=a^{\prime}\left(u^{\prime}\right)\left(b^{\prime}\left(u^{\prime}\right)\right)
$$

Remark (Distributivity of $a$ ). The role of ' $a$ ' in this combinator is distributive, in this sense that $\sqrt{67}$

$$
\forall a, b, u \cdot \bar{S}(a)(b)(u)=\bigcup_{x \in a} \bar{S}(\{x\})(b)(u)
$$

[^166]Therefore, we consider the existence of conditions (2a) - (2e) for $a=\{s\}$.

In other words, a proof that $(1 a-1 c)$ implies $(2 a-2 e)$ reduces to a proof that, for an arbitrary set $s$, given any sets $b$ and $u$ such that:

$$
\begin{align*}
& \chi(s) \subseteq u \\
& \wedge \chi \alpha(s) \subseteq \alpha^{"}\{y \in b \mid \chi(y) \subseteq u\}
\end{align*}
$$

One obtains a set $w$ such that:

$$
\begin{align*}
& \chi(w)=\{s\} \\
& \wedge \chi(s) \subseteq \chi \alpha^{2}(w) \subseteq u \\
& \wedge \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w) \\
& \wedge \chi \alpha(w) \subseteq b \\
& \wedge \alpha^{3}(w)=\alpha^{2}(s)
\end{align*}
$$

Now, for any arbitrary triple $(s, b, u)$, we derive sets $b_{s, u} \subseteq b$ and $u^{s, b} \subseteq u$ such that:

$$
\forall x . x \in\{s\}(u)(b(u)) \Longleftrightarrow x \in\{s\}\left(u^{s, b}\right)\left(b_{s, u}\left(u^{s, b}\right)\right)
$$

Definition 5.84 (Reduced forms of $b$ ). Given any three sets $b, s$ and $u$, we define the following subsets of $b$ :

$$
\begin{aligned}
& b_{s} \equiv\{y \in b \mid \alpha(y) \in \chi \alpha(s)\} \\
& b_{u} \equiv\{y \in b \mid \chi(y) \subseteq u\} \\
& b_{s, u} \equiv b_{s} \cap b_{u}=\{y \in b \mid \alpha(y) \in \chi \alpha(s) \wedge \chi(y) \subseteq u\}
\end{aligned}
$$

The following result is apparent, from the definition of $b_{s, u}$ :

$$
\forall s, b, u \cdot\{s\}(u)(b(u))=\{s\}(u)\left(b_{s, u}(u)\right)
$$

Equivalently:

$$
\forall s, b, u \cdot \alpha^{"} b_{s, u}=\chi \alpha(s) \Longleftrightarrow\{s\}(u)(b(u))=\alpha^{2}(s)
$$

Definition 5.85 (Reduced forms of $u$ ). Given any three sets $b, s, u$, we define the following subsets of $u$ :

$$
\begin{aligned}
& u^{s} \equiv \chi(s) \cap u=\{z \in u \mid z \in \chi(s)\} \\
& u^{b} \equiv \bigcup_{y \in b_{u}} \chi(y)=\{z \in u \mid \exists y \in b \cdot \chi(y) \subseteq u \wedge z \in \chi(y)\} \\
& u^{s, b} \equiv u^{s} \cup \bigcup_{y \in b_{s, u}} \chi(y)=\{z \in u \mid z \in \chi(s) \vee \exists y \in b \cdot \alpha(y) \in \chi \alpha(s) \wedge z \in \chi(y) \subseteq u\}
\end{aligned}
$$

As with $b_{s, u}$, it is apparent from the definition of $u^{s, b}$ :

$$
\forall s, b, u .\{s\}(u)(b(u))=\{s\}\left(u^{s, b}\right)\left(b\left(u^{s, b}\right)\right)
$$

Combining the reductions of $b$ and $u$, we obtain the following lemma:

Lemma 5.86. Given any three sets $s, b, u$ :

$$
\forall x . x \in\{s\}(u)(b(u)) \Longleftrightarrow x \in\{s\}\left(u^{s, b}\right)\left(b_{s, u}\left(u^{s, b}\right)\right)
$$

Proof. A proof of this result is equivalent to proving:

$$
\begin{aligned}
& \chi(s) \subseteq u \wedge \chi \alpha(s) \subseteq \alpha^{"}\{y \in b \mid \chi(y) \subseteq u\} \\
& \Longleftrightarrow \\
& \chi(s) \subseteq u^{s, b} \wedge \chi \alpha(s) \subseteq \alpha^{"}\left\{y \in b_{s, u} \mid \chi(y) \subseteq u^{s, b}\right\}
\end{aligned}
$$

One can see that $u^{s, b}$ and $b_{s, u}$ were constructed to be precisely the $\subseteq$-minimal subsets, satisfying this equivalence.

As $a$ is distributive in $a(u)(b(u))$, we obtain a corollary.

Corollary 5.87. Given any three sets $a, b, u$ :

$$
\forall x . x \in a(u)(b(u)) \Longleftrightarrow x \in \bigcup_{s \in a}\{s\}\left(u^{s, b}\right)\left(b_{s, u}\left(u^{s, b}\right)\right)
$$

Corollary 5.87 allows us to reduce $\chi$-continuity of $S$ to the following special case.

Corollary 5.88. The conditions entailing $\chi$-continuity of $S$ :

$$
\forall x . x \in a(u)(b(u)) \Longrightarrow x \in \bar{S}(a)(b)(u)
$$

can be reduced to a special case, which assumes three further conditions:

$$
\begin{aligned}
\exists s . a & =\{s\} \\
& \wedge \forall y \in b . \alpha(y) \in \chi \alpha(s) \wedge \chi(y) \subseteq u \\
& \wedge \forall z \in u . z \in \chi(s) \vee \exists y \in b . z \in \chi(y)
\end{aligned}
$$

Using Conditions ( $1 a^{\prime}$ ) and ( $1 b^{\prime}$ ), and the further conditions obtained from Corollary 5.88, we determine a set $w$, satisfying Conditions $\left(2 a^{\prime}\right)-\left(2 e^{\prime}\right)$, by constructing a tree diagram for which $w$ is a root. Two of the terminal nodes have already been defined:


It remains to determine the nodes $\chi \alpha(w)$ and $\chi \alpha^{2}(w)$, requiring minimal assumptions ${ }^{68}$ on the coding system and satisfying:

$$
\begin{align*}
& \chi(s) \subseteq \chi \alpha^{2}(w) \subseteq u \\
& \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w) \\
& \chi \alpha(w) \subseteq b
\end{align*}
$$

[^167]To satisfy these conditions, we consider two related ideas: $\chi$-connectivity and $\chi$-choice. The former arises from examination of the $\chi$-continuous fragment of $S$, the latter is a direct restriction of the axiom of choice in its multiplicative form $\sqrt{69}$ We first consider $\chi$-connectivity.

## $\chi$-Connectivity

$\chi \alpha(w)$ must satisfy Conditions $\left(2 c^{\prime}\right)$ and $\left(2 d^{\prime}\right)$. Together, these conditions imply, at least:

$$
\chi \alpha(w) \subseteq b \wedge \chi \alpha(s) \subseteq \alpha^{\prime \prime} \chi \alpha(w)
$$

Meanwhile, Condition ( $1 b^{\prime}$ ) implies $\chi \alpha(s) \subseteq \alpha^{\prime \prime} b$. Therefore, proving the existence of some set $\chi \alpha(w)$ satisfying Conditions ( $2 c^{\prime}$ ) and ( $2 d^{\prime}$ ) requires the existence of a set $z \in P_{\chi}(b)$ such that $\chi \alpha(s) \subseteq \alpha^{\prime \prime} z$.

It is not clear that a condition should exist, guaranteeing the general existence of such a set. However, Corollary 5.88 allows us to restrict our attention to the reduced case:

$$
\exists s \in a \cdot \chi \alpha(s)=\alpha^{"} b
$$

In such cases, $j^{\text {‘}} \alpha$ induces a $\chi \alpha(s)$-indexed partition of $b$ :

$$
\coprod_{x \in \chi \alpha(s)}\{y \in b \mid \alpha(y)=x\}
$$

Thus, the existence of $z \in P_{\chi}(b)$ such that $\chi \alpha(s) \subseteq \alpha^{\prime \prime} z$ is implied by the following condition.

Definition 5.89. A pre-combinatory algebra $(\alpha, \chi, \mathcal{A})$ is said to $\chi$-connected if, given any set $b$, any $\chi$-indexed partition $\coprod_{x} b_{x}$ of $b$ is "connected" by a set $y \in P_{\chi}{ }^{〔} b \sqrt{70}$

$$
\exists y \in P_{\chi}(b) . \forall b_{x} . y \cap b_{x} \neq \emptyset
$$

In addition to $\chi$-connectivity, we consider what is, in effect, a closure condition on our coding operation. $\chi$ is said to be closed when it is closed under sum-sets.

[^168]Definition 5.90. A coding $\chi$ is closed if, given any set $x$ and any $\chi(x)$-indexed family of coded sets:

$$
\exists y \cdot \chi(y)=\bigcup_{z \in \chi(x)} \chi(z)
$$

Remark (Closure, Choice, Connectivity). In the presence of an axiom that says something like: "given any $\chi$-indexed family of sets $\left(A_{z}\right)_{z \in \chi(x)}$, there exists a choice function between the indexing set and $\left(P_{\chi}\left(A_{z}\right)\right)_{z \in \chi(x)}$, mapping each $z \in \chi(x)$ to some coded subset of $A_{z}$," Definition 5.90 would imply Definition 5.89 .

Theorem 5.91. For any pre-combinatory algebra $(\alpha, \chi, \mathcal{A})$ the following conditions imply $\chi$-continuity of $S{ }^{711}$

1. $\forall s . \gamma(\{s\}) \neq \emptyset$
2. $(\alpha, \chi, \mathcal{A})$ is $\chi$-connected (Definition 5.89).
3. Given any sets $s$ and $t$ :

$$
\exists z \cdot \chi(z)=\chi(s) \cup \chi(t)
$$

4. The coding operation $\chi$ is closed (Definition 5.90).

Proof. As stated above, given any three sets $s, b$ and $u$, we seek to prove the existence of a set $w$, satisfying Conditions $\left(2 a^{\prime}\right)-\left(2 e^{\prime}\right)$. We determine $w$ by decorating the terminal nodes of its coding tree: $\chi(w), \chi \alpha(w), \chi \alpha^{2}(w)$, and $\alpha^{3}(w)$.

We have already determined canonical decorations for $\alpha^{3}(w)$ and $\chi(w)$, the existence of which follow from Definition 5.76 (Axiom 1) and Condition 1 of the current theorem:

$$
\begin{align*}
& \alpha^{3}(w)=\alpha^{2}(s)  \tag{Definition5.76,Axiom1}\\
& \chi(w)=\{s\}
\end{align*}
$$

(Condition 1)

It remains to determine the conditions for the existence of canonical $\chi \alpha(w)$ and $\chi \alpha^{2}(w)$, satisfying $\left(2 b^{\prime}\right)-\left(2 d^{\prime}\right)$.

[^169]$\chi \alpha(w)$ must satisfy:
\[

$$
\begin{align*}
& \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w) \\
& \chi \alpha(w) \subseteq b
\end{align*}
$$
\]

By Corollary 5.88 and Condition 2 of our hypothesis (that $(\alpha, \chi, \mathcal{A})$ is $\chi$-connected):

$$
\exists v \in P_{\chi}(b) . \forall z \in \chi \alpha(s) \cdot v \cap \alpha^{-1}(z) \upharpoonright b \neq \emptyset
$$

$\chi \alpha(w)=v$ satisfies Condition (2d $)$ and part of Condition (2c').
It remains to determine $\chi \alpha^{2}(w)$ satisfying:

$$
\begin{align*}
& \chi(s) \subseteq \chi \alpha^{2}(w) \subseteq u \\
& \forall t \in \chi \alpha(s) \cdot \exists y \in \chi \alpha(w) \cdot \alpha(y)=t \wedge \chi(y) \subseteq \chi \alpha^{2}(w)
\end{align*}
$$

As $v \in P_{\chi}(b)$, it clearly follows:

$$
\bigcup_{y \in v} \chi(y) \in P_{\chi}(u)
$$

Therefore, by Condition 4 of our hypothesis,

$$
\exists t \cdot \chi(t)=\bigcup_{y \in v} \chi(y)
$$

Furthermore, we know:

$$
\{s\}(b)(u) \neq \emptyset \Longrightarrow \chi(s) \in P_{\chi}(u)
$$

Condition 3 of our hypothesis then implies:

$$
\exists r \cdot \chi(r)=\chi(s) \cup \bigcup_{y \in v} \chi(y)
$$

It follows that $\chi \alpha^{2}(w)=\chi(s) \cup \bigcup_{y \in v} \chi(y)$ satisfies Condition (2b') and, by construction, $\chi \alpha(w)$ and $\chi \alpha^{2}(w)$ satisfy Condition (2c').

We obtain a set $w$, satisfying Conditions $\left(2 a^{\prime}\right)-\left(2 e^{\prime}\right)$, classified by the coding tree:


## $\chi$-Choice

We now consider the relationship between $\chi$-continuity of $S$ and a restricted choice principle, we refer to as $\chi$-choice.

Definition 5.92 ( $\chi$-choice). Given a coding operation $\chi$, for a given model $\mathcal{M}$ of set theory, we say $\mathcal{M}$ has satisfies the axiom of $\chi$-choice if, given any $\chi(x)$-indexed partition $\coprod_{z \in \chi(x)} b_{z}$ of a set $b$, there exists a set $v$ such that $v \cap b_{z}$ is a singleton set, for each component $b_{z}$ of the partition.

Remark (Choice vs. Connectivity). It is not true, in general, that either $\chi$-connectivity or $\chi$-choice entail the other. Consider an arbitrary $\chi(x)$-indexed partition of a set $b$ :

$$
b=\coprod_{z \in \chi(x)} b_{z}
$$

We obtain the following implications:

$$
\begin{aligned}
& \chi \text {-choice } \Longrightarrow \exists v \in P(b) . \forall z \in \chi(x) \cdot \exists y_{x} \cdot v \cap b_{z}=\left\{y_{x}\right\} \\
& \chi \text {-connectivity } \Longrightarrow \exists w \in P_{\chi}(b) . \forall z \in \chi(x) \cdot v \cap b_{z} \neq \emptyset
\end{aligned}
$$

While it is the case that $\chi$-choice implies $v \cong \chi(x)$, we have made no assumptions on coded sets, regarding size. In other words, $v \cong \chi(x)$ does not necessarily imply $\exists y \cdot \chi(y)=v$. Thus, $\chi$-choice does not imply $\chi$-connectivity.

Likewise, $\chi$-connectivity does not necessarily imply $\chi$-choice. The existence of some set $w \in P_{\chi}(b)$, such that has a non-empty intersection with each member of $\coprod_{z \in \chi(x)} b_{z}$, does not imply the existence of some other set $w^{\prime} \in P_{\chi}(b)$, such that $\forall z \in \chi(x) \cdot \exists y_{z} \cdot w^{\prime} \cap b_{z}=$ $\left\{y_{z}\right\}$.

In this way, both directions of the implication correspond to implicit size assumptions, regarding $\chi$. The former holds if coded sets are determined by cardinal size (e.g. 58]). The latter holds if $\chi$ is closed under subsets:

$$
\forall w \cdot \exists w^{\prime} \cdot \chi\left(w^{\prime}\right)=w \Longrightarrow P(w)=P_{\chi}(w)
$$

While neither $\chi$-choice nor $\chi$-connectivity are dependent on the other, generally, in the context of Theorem 5.91 we obtain an implication in one direction.

Lemma 5.93. If a coding operation $\chi$ is closed (Definition 5.90) and distributive (i.e. $\forall s \exists z \cdot \chi(z)=\{s\}), \chi$-choice implies $\chi$-connectivity ${ }^{[72}$

Proof. Given a $\chi$-indexed partition $\coprod_{z \in \chi(x)} b_{z}$ of a set $b$, $\chi$-choice implies the existence of a set $v$ such that each $v \cap b_{z}$ is a singleton set, for each component of the partition. We denote the corresponding singleton as $\left\{v_{z}\right\}=v \cap b_{z}$. As $\chi$ is distributive, there exists some set $y_{z}$ such that $\chi\left(y_{z}\right)=\left\{v_{z}\right\}$, for each $b_{z}$. As $\chi$ is closed, there exists a set $y$ such that:

$$
\chi(y)=\bigcup_{\chi(x)}\left\{v_{z}\right\}=\bigcup_{\chi(x)} \chi\left(y_{z}\right)
$$

Therefore, any $\chi$-indexed partition is "connected" by a coded set of the form $\chi(y)$, as constructed above.

As a direct corollary to Lemma 5.93 and Theorem 5.91, we obtain the result:

[^170]$$
\chi \text {-choice }+' \chi \text { is closed' }=\chi \text {-connectivity }
$$

Corollary 5.94. For any pre-combinatory algebra $(\alpha, \chi, \mathcal{A})$ the following conditions imply $\chi$-continuity of $S$ :

1. $\forall s . \gamma(\{s\}) \neq \emptyset$
2. $(\alpha, \chi, \mathcal{A})$ satisfies $\chi$-choice.
3. Given any sets $s$ and $t$ :

$$
\exists z \cdot \chi(z)=\chi(s) \cup \chi(t)
$$

4. The coding operation $\chi$ is closed.

In some ways, Corollary 5.94 is more familiar ${ }^{73}$ than Theorem 5.91 - intuitively, the principle $\chi$-choice is more "set-theoretic" than $\chi$-connected. On the other hand, the conditions for Theorem 5.91 are more general.

## Coding as Smallness

So far, we have made every effort not to assume that "coding" is determined by "size.," ${ }^{74}$ In comparing $\chi$-choice and $\chi$-connectivity, we considered two possibilities that we will now combine. The first size-related property is that $\chi$ is closed under subsets:

$$
\forall x \cdot(\exists z \cdot \chi(z)=x) \Longrightarrow(\forall y \cdot y \subseteq x \Longrightarrow \exists w \cdot \chi(w)=y)
$$

The second property can be stated: given a cardinal equivalence class $\kappa$, either every set $x \in \kappa$ is coded, or no set $x \in \kappa$ is coded:

$$
\forall \kappa \in N C \cdot(\exists x \in \kappa \cdot \exists y \cdot \chi(y)=x) \Longrightarrow(\forall z \in \kappa \cdot \exists w \cdot \chi(w)=z)
$$

Taken together, these imply $\sqrt{75}$

$$
\forall \kappa \in N C .(\exists x \in \kappa \cdot \exists y \cdot \chi(y)=x) \Longrightarrow(\forall z \cdot|z| \prec \kappa \Longrightarrow \exists w \cdot \chi(w)=z)
$$

[^171]The coding in [57, 58], where $\chi$-continuity corresponds to "finite character" clearly satisfies this condition. In particular, we can associate a coding satisfying this condition with a maximal cardinal $\kappa$, such that any set $x$ of cardinality less than $\kappa$ is coded. The principle of countable continuity, induced by application for NF with Quine Sequences, requires a non-trivial choice condition for $\chi$-continuity of $S$.

Proposition 5.95. In a model of NF + Countable Choice, $S$ is $\omega$-continuous when application is defined as app $\boldsymbol{a}_{\omega}$ (Definition 5.63).

Proof. This is just a special case of Lemma 5.94 .

Both in the context of these further assumptions and in general, it would be interesting if we had a result saying that $\chi$-choice (or $\chi$-connectivity) was necessary for the formation of $S$ but, at the moment, such a result is not apparent to the author ${ }^{76}$

[^172]
## Bibliography

[1] T. Altenkirch, J. Chapman, and T. Uustalu. Monads Need Not Be Endofunctors. Logical Methods in Computer Science, pages 1-40, 2015.
[2] J. Bènabou. Problèmes dan les topos. Technical Report Rapport 34, Univ. Cath. de Louvain, 1973.
[3] J. Bènabou. Fibered Categories and the Foundations of Naive Category Theory. The Journal of Symbolic Logic, pages 10-37, 1985.
[4] J. Bènabou. Distributors at Work (Lecture Notes), 2000.
[5] F. Borceux. Handbook of Categorical Algebra. Cambridge University Press, 1994.
[6] S. Eilenberg and S. Mac Lane. General Theory of Natural Equivalences. Transactions of the American Mathematical Society, pages 231-294, 1946.
[7] S. Eilenberg and J.C. Moore. Adjoint Functors and Triples. Illinois Journal of Mathematics, pages 381-398, 1965.
[8] M. Fiore, N. Gambino, M. Hyland, and G. Winskel. Relative Pseudomonads, Kleisli Bicategories, and Substitution Monoidal Structures. Sel. Math. New Series, pages 2791-2830, 2018.
[9] T.E. Forster. Set Theory With A Universal Set. Oxford University Press, 2nd edition, 1992.
[10] T.E. Forster. Why Set Theory Without Foundation? Journal of Logic and Computation, 4:333-335, 1994.
[11] T.E. Forster. Implementing Mathematical Objects in Set Theory. Logique et Analyse, 2007.
[12] T.E. Forster. Our Exagimation round his Factification: a Backgrounder for Randall Holmes' Proof of the Consistency of Quine's NF, 2017.
[13] T.E. Forster and R.M. Kaye. End-extensions Preserving Power Set. Journal of Symbolic Logic, pages 323-328, 1991.
[14] T.E. Forster, A. Lewicki, and A. Vidrine. Category Theory with Stratified Set Theory, 2019. arXiv:1911.04704.
[15] J.D. Gergonne. Considerations Philosophiques sur les elements de la science de l'etendue. Ann. Math. Pures Appl., pages 209-231, 2007.
[16] T. Hailperin. A Set of Axioms for Logic. Journal of Symbolic Logic, pages 1-19, 1944.
[17] C.W. Henson. Permutation Methods Applied to NF. Journal of Symbolic Logic, pages 69-76, 1973.
[18] C.W. Henson. Type-Raising Operations on Cardinal and Ordinal Numbers in Quine's "New Foundations". Journal of Symbolic Logic, pages 59-68, 1973.
[19] M.R. Holmes. Elementary Set Theory With A Universal Set. Cahiers Du Logique, 1992.
[20] M.R. Holmes. The Equivalence of NF-Style Set Theories with "Tangled" Theories; The Construction of $\omega$ Models of Predicative NF (and more). Journal of Symbolic Logic, pages 178-190, 1995.
[21] R.B. Jensen. On the Consistency of a Slight(?) Modification of Quine's NF. Synthese, pages 250-263, 1969.
[22] P.T. Johnstone. Topos Theory. Academic Press, 1977.
[23] P.T. Johnstone. Sketches of an Elephant: A Topos Theory Compendium. Oxford University Press, 2002.
[24] A. Joyal and I. Moerdijk. Algebraic Set Theory. Cambridge University Press, 1995.
[25] A. Kock. Limit Monads In Categories. PhD thesis, University of Chicago, 1967.
[26] A. Kock. Monads For Which Structures Are Adjoint to Units. Journal of Pure and Applied Algebra, pages 41-59, 1995.
[27] J. Lambek. Subequalizers. Canadian Math. Bulletin, pages 337-349, 1970.
[28] J. Lambek and P.J. Scott. Intuitionist Type Theory and the Free Topos. Journal of Pure and Applied Algebra, pages 215-257, 1980.
[29] F.W. Lawvere. An Elementary Theory of the Category of Sets. Proceedings of the National Academy of Science of the U.S.A, pages 1506-1511, 1965.
[30] F.W. Lawvere. Diagonal Arguments and Cartesian Closed Categories. Reprints in Theory and Applications of Categories, pages 1-13, 2006.
[31] T. Leinster. The Yoneda Lemma: What's it all about?, 2000.
[32] S. Mac Lane. Mathematics Form and Function. Springer-Verlag, 1986.
[33] S. Mac Lane. Categories for the Working Mathematician. Springer-Verlag, 2nd edition, 1998.
[34] M. Makkai. Avoiding the Axiom of Choice In General Category Theory. Journal of Pure and Applied Algebra, pages 109-173, 1996.
[35] E. Manes. Monads of Sets. In Handbook of Algebra, pages 65-154. North-Holland Publishing Company, 2003.
[36] F. Marmolejo and R. Wood. Doctrines Whose Structure Forms A Fully Faithful Adjoint String. Theory Appl. Categ., pages 22-42, 1997.
[37] F. Marmolejo and R. Wood. Kan Extensions and Lax Idempotent Pseudomonads. Theory Appl. Categ., pages 1-29, 2012.
[38] F. Marmolejo and R. Wood. No-Iteration Pseudomonads. Theory Appl. Categ., pages 371-402, 2013.
[39] C. Maurer. Universes in Topoi. In Model Theory and Topoi, pages 284-296. Springer, 1975.
[40] C. McLarty. Failure of Cartesian Closedness in NF. Journal of Symbolic Logic, pages 555-556, 1992.
[41] A.R. Meyer. What is a model of the lambda calculus? Information and Control, pages 87-122, 1982.
[42] A. Nerode. Some Stone Spaces and Recursion Theory. Duke Mathematical Journal, pages 397-406, 1959.
[43] D. Pavlovic. On the Structure of Paradoxes. Archives for Mathematical Logic, pages 397-406, 1992.
[44] G.D. Plotkin. A set-theoretical definition of application. Technical Report Memo. MIP-R-95, School of Artificial Intelligence, University of Edinburgh, 1972.
[45] W.V. Quine. New Foundations for Mathematical Logic. American Mathematical Monthly, pages 70-80, 1937.
[46] W.V. Quine. Mathematical Logic. Harvard University Press, 1940.
[47] J.C. Reynolds. Polymorphism is not Set-theoretic. In Semantics of Data Types, pages 223-265. Springer, 1984.
[48] J.B. Rosser. The Burali-Forti Paradox. Journal of Symbolic Logic, pages 1-17, 1942.
[49] J.B. Rosser. The Axiom of Infinity In Quine's New Foundations. Journal of Symbolic Logic, pages 238-242, 1952.
[50] J.B. Rosser. Logic For Mathematicians. Dover Publications Inc., 2nd edition, 1978.
[51] B. Russell. Mathematical Logic As Based on The Theory of Types. American Journal of Mathematics, 1908.
[52] B. Russell. Introduction to Mathematical Philosophy. Routledge, 1919.
[53] A. Salibra. Lambda calculus: Models and theories. In Proceedings of the Third AMAST Workshop on Algebraic Methods in Language Processing. University of Twente, 2003.
[54] D. Scott. Quine's Individuals. In Logic, Methodology and Philosophy of Science, pages 111-115. Stanford University Press, 1962.
[55] D. Scott. Continuous Lattices. In Toposes, Algebraic Geometry and Logic, Lecture Notes in Mathematics, vol 274., pages 97-136. Springer-Verlag, 1972.
[56] D. Scott. Models for Various Type-Free Calculi. In Proc. IVth Internat. Cong. for Logic, Methodology and the Philosophy of Science., pages 157-187. North-Holland, 1973.
[57] D. Scott. Datatypes As Lattices. SIAM Journal of Mathematics, pages 57-66, 1976.
[58] D. Scott. Lambda Calculus: Some Models, Some Philosophy. In The Kleene Symposium, pages 223-265. North-Holland Publishing Company, 1980.
[59] D. Scott and D. McCarty. Reconsidering Ordered Pairs. The Bulletin of Symbolic Logic, pages 379-397, 2008.
[60] A.K. Simpson. Elementary Axioms for Categories of Classes. In Proceedings of the 14th Annual IEEE Symposium on Logic in Computer Science, pages 77-85. 1999.
[61] E. Specker. The Axiom of Choice in Quine's New Foundations. Proceedings of the National Academy of Sciences of the USA, pages 972-975, 1953.
[62] E. Specker. Dualität. Dialectica, pages 451-465, 1958.
[63] E. Specker. Typical Ambiguity. In Logic, Methodology and Philosophy of Science, pages 116-123. Stanford University Press, 1962.
[64] R. Street. The Formal Theory of Monads. Journal of Pure and Applied Algebra 2, pages 149-168, 1972.
[65] R. Street and R. Walters. Yoneda Structures on 2-Categories. Journal of Algebra, pages 350-379, 1978.
[66] T. Streicher. Fibred Categories à la Jean Bènabou, 2018. arXiv:1801.02927.
[67] P. Taylor. Practical Foundations of Mathematics. Cambridge University Press, 1999.
[68] F. Ulmer. Properties of Dense and Relative Adjoint Functors. Journal of Algebra, pages 77-95, 1968.
[69] C. Walker. Yoneda Structures and KZ Doctrines, 2017. arXiv:1703.08693.
[70] H. Wang. A Formal System of Logic. The Journal of Symbolic Logic, pages 25-32, 1950.
[71] H. Wang. Negative Types. Mind, pages 366-8, 1952.


[^0]:    ${ }^{1}$ Here we mean "small" in the category theoretic sense (i.e. internal). In NF, category theoretic smallness does not correspond to set theoretic smallness (i.e. cardinality). The appropriate notion of (set theoretic) smallness for NF is strongly cantorian. When refering to a category that is small in the sense of NF, we refer to it as NF-small. e.g. A category with strongly cantorian hom-sets is referred to as locally $N F$-small.

[^1]:    ${ }^{2}$ In most cases, the "stratified" analogues of structures that form adjunctions in Set possess stronger symmetry than is typical of relative adjunctions. Such structures motivate the abstract definition of a symmetric lift.
    ${ }^{3}$ A central example for NF is the implementation of natural numbers as Frege rather than Von Neumann naturals [11].

    4"Degenerate" in the sense that objects are enriched over $\{\top, \perp\}$, rather than Set.

[^2]:    ${ }^{5}$ Relative adjoints were first introduced by Ulmer [68]. The study of relative algebra was extended to (relative) monads in [1 and, more recently, to pseudomonads in [8].
    ${ }^{6}$ The history of TST (typed set theory) is somewhat unclear. The original idea of typing to avoid paradoxes is due to Russell [51. TST presents an evolution of Russell's ideas, the first formal presentation of which was written down by Tarski.
    ${ }^{7}$ Two important references for NF are [9] and 50.
    ${ }^{8}$ See Section 1.4

[^3]:    ${ }^{9}$ Our interest in this structure is not solely its generality. What is of primary interest to the author is that starting with examples of symmetric lifts in NF, one is naturally led to identify general category theoretic structures, such as Yoneda Extensions, that satisfy the same abstract properties.
    ${ }^{10}$ Chapter 3 is largely contained in 14. Any mistakes should be considered my own, and not those of Forster and Vidrine.
    ${ }^{11} S C U$ is the axiom stating: the sum-set taken over any family of strongly cantorian sets, indexed by a strongly cantorian set, is strongly cantorian.
    ${ }^{12}$ Throughout the literature on NF, a number of extensions have been considered in relation to the strongly cantorian sets - the Axiom of Counting being the most prominent example [50.

[^4]:    ${ }^{13}$ There is reason to expand this analysis to fibred category theory over $\mathcal{N}$ but, with some exceptions, we focus primarily on the internal category theory.
    ${ }^{14}$ See Section 4.5.
    ${ }^{15}$ Introduced in [8].

[^5]:    ${ }^{16}$ The classification does hold for the "free" relative pseudoalgebras (i.e. the relative Kleisli algebras) [8]. But the proof does not extend to the broader category of (potentially non-free) relative EM-algebras (see Lemma 4.44).
    ${ }^{17}$ There are caveats. The failure of cartesian closure rules out any straightforward application operation and the existence of a universal set and complementation raises concerns regarding consistency.

[^6]:    ${ }^{18}$ For the record: Sally says to her brother, "I'm older than you, so I get the larger half of the brownie." Explain why this statement is incorrect. (Hint: It isn't a matter of justice.)

[^7]:    ${ }^{1}$ Examples include: [29, [32], [24], [67, Chapter 9].
    ${ }^{2}$ Although the strength required in the base theory, to develop a reasonable theory of categories, is less than one might think 3].
    ${ }^{3}$ In the category theoretic sense (i.e. internal).

[^8]:    ${ }^{4}$ See Theorem 4.48
    ${ }^{5}$ Relative adjunctions were introduced by Ulmer and, more recently, extended by Alternkirch and Hyland et al [68, 1, 8].
    ${ }^{6}$ See Section 2.2.1 and Definition 2.12 .

[^9]:    ${ }^{7}$ From the standpoint of NF, this gives a fuller semantic picture of the $T$-functor (i.e. $x \mapsto \iota " x$ ).
    ${ }^{8}$ By "relative KZ-pseudomonad," we mean a lax idempotent relative pseudomonad, in the sense of [8].
    ${ }^{9}$ The category of strongly cantorian sets is locally small, but if it were to exist in $\operatorname{cat}(\mathcal{N})$, we would obtain a form of the Burali-Forti Paradox (Proposition 3.78).
    ${ }^{10}$ The forerunner of TST was the type theory of Russell [51. While the exact history is somewhat opaque, it seems Tarski was the first to introduce what we would recognize as TST today.

[^10]:    ${ }^{11}$ This assumes implementation of type-level ordered pairs, which we expand upon later.

[^11]:    ${ }^{12}$ Fraenkel-Mostowski models can be used to construct a model of TST where Choice fails.
    ${ }^{13}$ This anticipates the occurrence of relative adjunctions in the category theory of NF.

[^12]:    ${ }^{14}$ Any sensible definition of ordered pair $\langle x, y\rangle$ must be defined with $x$ and $y$ occuring at the same level.
    ${ }^{15} \iota$ " $x_{i}$ indicates the set $\left\{\left\{z_{i-1}\right\} \mid z_{i-1} \in{ }_{i-1} x_{i}\right\}$.

[^13]:    ${ }^{16}$ If NF had a sensible notion of rank, this result would translate smoothly. But NF is not a well founded theory of sets.

[^14]:    ${ }^{17}$ The author is grateful to Randall Holmes for bringing the full importance of this classification to his attention. As far as the author is aware, Holmes is the first to introducelateral functions and should be credited accordingly.

[^15]:    ${ }^{18}$ This section is truly a "preview" in the sense that it assumes some knowledge of internal category theory (See Section 4.1). The reader is welcome to skip directly to Section 1.3

[^16]:    ${ }^{19}$ See Proposition 2.25
    ${ }^{20} \mathrm{We}$ assume type-level pairing.
    ${ }^{21} \mathrm{An}$ more thorough introduction to internal category theory can be found in Section 4.1 .

[^17]:    ${ }^{22}$ The action of $\widehat{(-)}$ on internal functors $J: \mathbb{C} \rightarrow \mathbb{D}$ in $C a t_{i}$ (i.e. pairs of maps $\left\langle J_{0}: C_{0} \rightarrow D_{0}, J_{1}\right.$ : $\left.C_{1} \rightarrow D_{1}\right\rangle$, satisfying $d_{i} \circ J_{1}=J_{0} \circ d_{i}$ ) is induced by pullback.
    ${ }^{23}$ The implications for higher category theory could be interesting, but the author has not moved beyond conjecture.

[^18]:    ${ }^{24}$ Importantly, while $x \cup\{x\}$ is a set of NF, one cannot "internalize" substitution. In other words, the graph of the operation $x \mapsto x \cup\{x\}$ is unstratified.

[^19]:    ${ }^{25}$ The version of this used above, $\{1\} \in 1$, corresponds to a slightly different statement of this property: $\forall n \in N .\{m \mid 0<m \leq n\} \in n$.

[^20]:    ${ }^{26}$ Implementation of homogeneous pairs, by Quine Pairing, is reviewed in Section 1.3.2
    ${ }^{27}$ For applications of NF's recursion theorem, see Section 5.4 .

[^21]:    ${ }^{28}$ There are many treatments of this in the literature; ours closely follows [50, Chapter 13].
    ${ }^{29} N F+$ AxCount $\vdash \operatorname{Con}(N F)$. .9, Theorem 2.3.8]

[^22]:    ${ }^{30}$ This can be extended to any $n$-length sequence. See Section 5.4

[^23]:    ${ }^{31}$ One might think ' $a \in f(a)$ ' is weakly stratified and, therefore, defines a set. If one writes the set abstract in primitive notation, however, one can see ' $a$ ' is bound. From this we extract a general rule: in set abstracts $\{z \mid \Phi(\vec{x}, z)\}$, the eigenvariable $z$ is considered bound in $\Phi$.

[^24]:    ${ }^{32}$ In both cases, the definition of $T$ is independent of the choice of a particular member of the equivalence class.

[^25]:    ${ }^{33} T^{-1}$ is the obvious (partially defined) operation, inverse to $T$.
    ${ }^{34}$ See 61. Roughly: As concrete natural numbers, $T(1)=1$ and $T(2)=2$. Also, $T(n+m)=$ $T(n)+T(m)$. So we obtain $n=T n+1=T(n+1)$ or $n=T n+2=T(n+2)$, both of which contradict the result that, given a finite $n$, only one of the following three cases can be true: 1) There exists some $m$ such that $n=m+m+m ; 2$ ) There exists some $p$ such that $n=p+p+p+1 ; 3$ ) There exists some $q$ such that $n=q+q+q+2$.

[^26]:    ${ }^{35}$ This is said to be "abstract" as it does not depend on any specific axiom or implementation in set theory 48.

[^27]:    ${ }^{36}$ See Theorem 3.64

[^28]:    37 "Urelemente" are more commonly referred to - at least in the author's experience - as atoms.
    ${ }^{38}$ It is obvious, but important to note, NFU is still a one-sorted theory. The predicate set simply corresponds to the condition expressed (in $\mathcal{L}_{\text {Set }}$ ) by the second axiom.
    ${ }^{39}$ The differences are, to some extent, captured by the idea that things tend to go "wrong" for NF because of the cardinal arithmetic of particularly large cardinals (e.g. the failure of choice, Theorem 1.33). The existence of atoms head off a number of these issues - consider why, in the context of NFU, $V$ may well be smaller than $P V$ - but discussion of this is beyond the scope of the current introduction. For an excellent exposition on why NFU is, in some sense, the more intuitive version of NF, see [20]. For the only introductory text on NFU (of which the author is aware) see 19 .

[^29]:    ${ }^{40}$ See Chapter 3. $\mathbf{C E}$ is required for the general existence of coequalizers and (internal) colimits. $\mathbf{S C U}$ is required to prove the class of strongly cantorian maps is closed (Definition 3.69), and enables a much tighter relationship between modified-dependent products of "small" display maps.

[^30]:    ${ }^{41}$ Notice that, over the collection of all objects, this is nothing more than the axiom of separation applied to the set $V$ of all sets.
    ${ }^{42}$ The reason for superscripts ranging over $N$ is obvious (to keep track of stratification), but this is purely convention - the theory is one-sorted.

[^31]:    ${ }^{43}$ The author should acknowledge Randall Holmes for bringing his attention to the definition used here.
    ${ }^{44}$ Equally, $F$ is $-n$-lateral if and only if $\operatorname{set}\left(\left\{\left\langle x, \iota^{n ‘} F^{\iota} x\right\rangle \mid x \in V\right\}\right)$.

[^32]:    ${ }^{45}$ Note, $\mathcal{N}$ is also an internal category in itself, but the internal functors of $\operatorname{cat}(\mathcal{N})$ form a proper subclass of those in the category of ML classes.
    ${ }^{46}$ A reader interested in NF only insofar as is required to understand the category theory of NF can safely skip this appendix.
    ${ }^{47}$ Specker used the system TZT. TZT is equivalent to TST, but levels range over all integers [71.
    ${ }^{48}$ The version of [62] familiar to the author is an annotated translation, due to Forster.

[^33]:    ${ }^{49}$ Holmes has produced the first credible proof of the consistency of NF, but it remains unpublished. The proof is based on earlier work in [20] and the best (and only) outline of the current version is [12].

[^34]:    ${ }^{50}$ There is an obvious symmetry in the two primitive predicates of projective geometry, but any permutation of a given language, respecting logical operations, gives rise to a valid form of "duality." For example, the implict asymmetry of the set membership relation does not invalidate the permutation of $\mathcal{L}_{T Z T}$ given above.
    ${ }^{51}$ In the case of $\mathcal{M} \models \mathcal{G}$, a correlation is given by a map exchanging points and lines, which preserves incidence.

[^35]:    ${ }^{52}$ The more general version of this lemma, replacing "automorphism" with "endomorphism" is given in 63].
    ${ }^{53}$ An important point is that the proof of $\neg A C$ in NF carries over to $T Z T$. Thus, choice fails in any model of $T Z T+A m b$. As Forster points out, this somewhat eliminates the hope that any elementary method could be used to form such a model.

[^36]:    ${ }^{1}$ A classical adjunction is just a relative adjunction, mediated by an identity functor.
    ${ }^{2}$ This is made formal in 65].

[^37]:    ${ }^{3}$ The image of a map under the natural isomorphism is referred to as its adjoint transpose.

[^38]:    ${ }^{4}$ We use the phrase "syntactic choice" as it arises from the comprehension/separation aspect of set theory.

[^39]:    ${ }^{5}$ What is interesting about this particular construction is not confined to the fact that it furnishes an example of a symmetric lift in a context more general than NF. By first examining the category of NF sets in isolation, one is naturally led to expect structures in $\mathcal{N}$ and the abstract (higher) category theoretic structure of a Yoneda Extension are specific cases of a common general structure (i.e. symmetric lifts).

[^40]:    ${ }^{6}$ It must be noted that, despite the (hopefully) compelling nature of this claim, outside of Proposition 2.17 the author's investigation of symmetric lifts outside NF remains in its early stages. Definition 2.12 implies that the "relative" functors, $J_{0}$ and $J_{1}$, of a symmetric lift are full and faithful for a certain class of morphisms. One can also see a likely connection between locally fully faithful lax idempotent relative pseudomonads (combining the relative pseudomonads of [8] and locally fully faithful KZ-Doctrines of [69) and some form of relative Yoneda Structure. Effectively, this would be a relative version of a result by Walker 69.
    ${ }^{7}$ The use of the term "symmetric lift" refers to Street's study of monads/adjunctions in general 2-categories, where relative adjoints correspond to absolute lifts 65].

[^41]:    ${ }^{8}$ A broader introduction to the Yoneda Lemma is given in Section 4.2 ,
    ${ }^{9}$ The latter relative adjunction is just a restriction of the standard adjunction $\widehat{F} \dashv F^{*}$.

[^42]:    ${ }^{10}$ Just as we could define $P f: P A \rightarrow P B$, for any function $f: A \rightarrow B$.
    ${ }^{11}$ From this we also obtain commutativity of the overall Yoneda Extension diagram: $\widehat{F} \circ \mathcal{Y}_{\mathcal{A}}=\mathcal{Y}_{\mathcal{B}} \circ F$.
    ${ }^{12}$ Notice that this coincides precisely with the (pointwise) construction of a left Kan Extension.

[^43]:    ${ }^{13} \gamma_{\tau}$ and $\gamma_{\theta}$ correspond to natural transformations $\tau: \mathcal{A}(-, A) \rightarrow \mathcal{B}(F-, B)$ and $\theta: \mathcal{A}\left(-, A^{\prime}\right) \rightarrow$ $\mathcal{B}(F-, B)$, respectively. $F(g) \circ-$ is the image $\mathcal{B}(-, F(g))$ of a natural transformation $g \circ-: \mathcal{A}(-, A) \rightarrow$ $\mathcal{A}\left(-, A^{\prime}\right)$, which is a morphism $g: \tau \rightarrow \theta$ in the comma category $\mathcal{Y}_{\mathcal{A}} \downarrow \mathcal{B}(F-, B)$.

[^44]:    ${ }^{14}$ We will frequently use the presentation of monads as extension systems, due to Manes 35].

[^45]:    ${ }^{15}$ If we do not have the latter, we can carry out an argument along the lines of McLarty or the general paradoxical structures of Lawvere and Pavlovic [30, 43].

[^46]:    ${ }^{16}$ In Chapter 5, we prove that a model of NF can be interpreted as the algebra of its distributive functions, where extensionality is exchanged for $\eta$-equivalence in the sense of [57].
    ${ }^{17} \mathrm{~A}$ good account of this can be found in (4).

[^47]:    ${ }^{18}$ See [8] for a general account of 2-categorical relative algebra.

[^48]:    ${ }^{19}$ We are getting something akin to a factorization system, where we are interested in the original transpose of maps $T C \rightarrow F D$ that factor through the class of maps in $\mathcal{C}$ defined by $T$-algebras.

[^49]:    ${ }^{1}$ As stated in the Declaration, Chapter 3is primarily the result of collaborative work with Forster and Vidrine [14]. Section 3.3 is a result of the author's independent work, as is the presentation of pseudo and locally modified-dependent products as symmetric lifts (a concept introduced in this thesis). But these results should not be considered as more than extensions/revisions (to the benefit or detriment) of work in 14. The writing of this chapter and the presentation of results reflects the decisions and work of the author. Any errors should be considered my own, and not those of Forster and Vidrine.

[^50]:    ${ }^{2}$ As $\mathcal{N}$ has a universal set, even $F a m \mathcal{C}$ is an internal category.
    ${ }^{3}$ Our use of the word "display" corresponds to the more formal use in 67.
    ${ }^{4} \mathrm{KF}$ was introduced in 13.

[^51]:    ${ }^{5}$ As opposed to restrictions one might observe on account of stratified comprehension.

[^52]:    ${ }^{6}$ As such, one may also wish to consider extension of $N F$ to include IO or $\mathbf{C E}$.
    ${ }^{7}$ We assume no sets beyond those that can be proven from the axioms of NF.

[^53]:    ${ }^{8}$ See Section 4.1
    ${ }^{9}$ One can even iterate this process to obtain: $\bigcup_{n \in N} \operatorname{cat}^{n}(\mathcal{N})$, by interpreting $n$-categories as $n+1$ categories, with trivial $n+1$-cells.
    ${ }^{10}$ See Theorem 3.78 .

[^54]:    ${ }^{11}$ Recall, a NF-small category is one whose collection of morphisms forms a strongly cantorian set. Thus, mirroring the classical definition of local smallness, a category is locally NF-small if each of its individual hom-sets is strongly cantorian.

[^55]:    ${ }^{12}$ We do note use the word "practical" casually. As the language of category theory (i.e. composition and identity) is homogeneous, any categorical structure defined diagramatically will correspond to stratified formula in $\mathcal{L}_{\text {Set }}$.
    ${ }^{13}$ Recall, definability by a stratified set abstract is a sufficient but not a necessary condition for the existence of a set.
    ${ }^{14} \mathrm{ML}$ is equiconsistent with NF [70].
    ${ }^{15}$ If the reader is comfortable thinking about external functors as simply objects of some generic meta-theorem (in which our model of NF exists), this section can safely be skipped. In particular, for

[^56]:    ${ }^{20}$ Practically speaking, we will not encounter any issues related to this problem within this thesis - the functor/diagram categories that interest us are all internal. Nevertheless, if we are helping ourselves to external functors (and natural transformations between them), we should have something to say about the category they might form.

[^57]:    ${ }^{21}$ For this reason, as discussed at the end of this section, there is little category theoretic structure that exists in the category of ML classes that would tempt us to work there, rather than in $\mathcal{N}$.

[^58]:    ${ }^{22}$ As mentioned above, the fact that ML is equiconsistent with NF makes it particularly useful, given our interest in developing the category theory of NF, while making as few assumptions as possible regarding our metatheory.

[^59]:    ${ }^{23}$ We could define an iterated union operation $\cup^{n}: V \rightarrow V$, but this could not be extended to morphisms in a manner that would be functorial.
    ${ }^{24}$ In addition, as mentioned immediately above, it cannot be extended to morphisms in a functorial manner.
    ${ }^{25}$ See Definition 1.7 .

[^60]:    ${ }^{26}$ As noted above, the limiting cones "lift" only up to isomorphism.

[^61]:    ${ }^{27}$ This is not a trivial difference. Cantor's Theorem proves there are cases where $|\mathcal{N}(1, X)| \lesseqgtr|X|$.

[^62]:    ${ }^{28}$ See Proposition 3.36

[^63]:    ${ }^{29}$ If one prefers to view a topos as sheaves over a generalized topological space, this intuition still holds - Set corresponds to the presheaf category of the trivial (i.e. single element) space.

    30 "Paradoxical structures" are due to Lawvere and Pavlovic [30, 43. The relevant example for NF is due to McLarty 40.
    ${ }^{31}$ Forster, Vidrine and the author emphasize this more general approach in [14.

[^64]:    ${ }^{32} T_{B}: \mathcal{C} / B \rightarrow \mathcal{C} / T B$ is the obvious functor induced by $T$.
    ${ }^{33}$ Nathan Bowler should be acknowledged for originally pointing out this feature.

[^65]:    ${ }^{34} \mathrm{We}$ can conceive of a more general idea of syntactic universal properties, as the existence of a univer-

[^66]:    sal property (a semantic concept, defined only up to isomorphism) for a (possibly proper) subcategory of diagrams of a given shape, witnessed by an instance of comprehension/separation.
    ${ }^{35}$ As noted previously, coeq : $\mathcal{C} \rightrightarrows \rightarrow \cdot \mathcal{C}$ is functorial.

[^67]:    ${ }^{36}$ Note that this is not as strong a requirement as saying that the correspondence between $C$ and $\Pi$ is a set.

[^68]:    ${ }^{37}$ Whats holds locally in $\mathrm{KF}(\mathrm{I})$ holds globally in $\mathrm{NF}-$ recall: $N F=K F+\exists y \forall x . x \in y$.

[^69]:    ${ }^{38}$ Notice, in this case, the relative adjoint determines a map other than the semantic coequalizer, which clearly exists in the trivial case (i.e. the identity morphism).

[^70]:    ${ }^{39}$ For more on this, see 31.

[^71]:    ${ }^{40}$ See 3$]$.

[^72]:    ${ }^{41}$ With replacement, this is also true if one is working within a small category in Set.
    ${ }^{42}$ One can refine $\prec$ to include uniqueness of factorizations.

[^73]:    ${ }^{43}$ As the language of category theory is homogeneous, standard categorical structures carry stratified definitions.
    ${ }^{44}$ Typically, as with the coequalizer $\Phi_{f, g}$, the issue is not that the set abstract of the universal object is unstratified, but that it is inhomogeneous - i.e. the graph(s) of the universal morphism(s) is unstratified.
    ${ }^{45}$ This is not dissimilar to the idea behind permutation models 54.

[^74]:    ${ }^{46}$ Again, where $h$ satisfies this property, but $j$ is not unique, we can exchange $h$ for $i m(h)$, and obtain a coequalizer.

[^75]:    ${ }^{47} \exists h . \hat{h}=\downarrow \Psi_{f, g}$ yields the homogeneous map $z \mapsto \cup h^{"}[z]_{\sim_{h}}$.

[^76]:    ${ }^{48}$ To get the strict equivalence, use $R U S C$.

[^77]:    ${ }^{49}$ We will use ' $A \Rightarrow B$ ' to denote "literal" function spaces in $\mathcal{N}$, and reserve ' $B^{A}$ ' to denote exponential objects.

[^78]:    ${ }^{50}$ The stratified analogue to evaluation (i.e the co-unit) has been observed, independently, by Morgan Taylor.

[^79]:    ${ }^{51}$ Proposition 3.54 appeared in our work in [14], for a generic SPE (Definition 3.18). We include the more general proof from the original proposition, but one can easily compare it to the specific case of $\mathcal{N}$. The existence of both universal transformations is a theorem of NF, but this is not obvious for certain (weaker) extensions of KF.

[^80]:    ${ }^{52}$ Nathan Bowler first observed that $\operatorname{Sub}(T A \times-)$ is representable in NF.

[^81]:    ${ }^{53}$ Not least because there are limitations on internal displays of indexed families.

[^82]:    ${ }^{55}$ Using the fact that concrete finite sets are strongly cantorian.

[^83]:    ${ }^{56}$ The axioms of a category of classes are only one aspect of the stronger requirements for a class category.
    ${ }^{57}$ This defines a class category in its strong form - many would only require the universe to be some object $U$ such that $P_{\mathcal{S}} U \rightharpoondown U$. A nice corollary of working in the stronger version is the existence of an internal model of $\mathcal{S}$, the full internal subcategory of "sets" in $\mathcal{C}$ [60. $V \in \mathcal{N}$ implies a similar result for all NF sets (see Theorem 4.48) but, actually, confounds the interpretation of strongly cantorian sets as "small" (see Proposition 3.78).

[^84]:    ${ }^{58}$ The above result does not require SCU , and would also hold for the finite sets in $K F+I n f$. We might accordingly expect to find a subtopos of any category behaving like a category of stratified sets (i.e. any SPE).

[^85]:    ${ }^{59}$ Here one might observe that, intuitively, any monomorphism is a small map, as its fibres are singletons. Externally, this is true. In the internal language of a category, however, we only have equality in the sense of diagonal maps (thus, we only have equality as can be defined by the unit of the binary instance of the product adjunction). Therefore, asserting that each monomorphism is small, in the setting of a category of classes, is actually stating a form of replacement.

[^86]:    ${ }^{60}$ This is discussed in Section 4.6 in connection with an idea of Taylor 67.

[^87]:    ${ }^{1}$ Sections 4.1 and 4.2, as well as Theorem 4.19 and Theorem 4.48 , are the result of collaborative work with Vidrine, which also appears in our joint work with Forster [14. The writing and presentation, in the present form, reflect the work and decisions of the author. As always, errors should be considered mine alone.
    ${ }^{2}$ Relative pseudomonads have recently been defined by Hyland et al 8.
    ${ }^{3}$ At times, we may refer to a KZ-doctrine as a KZ-monad, but we always mean the same, 2dimensional structure. Background material on KZ-doctrines can be found in [26, 36, 37].

[^88]:    ${ }^{4}$ No-iteration pseudomonads are, effectively, a higher dimensional version of Manes-style monads (38.

[^89]:    ${ }^{5}$ Top is also an example of a category where the axiom of choice fails (not all continuous surjections have continuous splittings).

[^90]:    ${ }^{6}$ Note: This embedding is full and faithful, as any natural transformation between internal functor must itself be internal.
    ${ }^{7}$ Recall, $S C$ denotes the full subcategory of strongly cantorian sets in NF.

[^91]:    ${ }^{8}$ For further introduction, one can follow almost directly from this section to the corresponding section in [22].

[^92]:    ${ }^{9}$ Of course, we are also unable to disprove such a claim.
    ${ }^{10}$ A Group is simply a category $\mathcal{G}$ with a single object and a map $\tau: \operatorname{Mor}(\mathcal{G}) \rightarrow \operatorname{Mor}(\mathcal{G})$, taking each morphism to its inverse.

[^93]:    ${ }^{11}$ The following observation has been made on multiple occasions, but we are working from [4], in particular. Here, Bènabou is interested in motivating Distributors as the appropriate generalization of relations between categories. One can pursue this as well - to the end of $\operatorname{Rel}(\mathcal{N})$.

[^94]:    ${ }^{12}$ In the sense that it preserves suprema.

[^95]:    ${ }^{13}$ As opposed to the trivial selection function in the case of CE.
    ${ }^{14}$ The "selection" function mentioned above, allowing one to subvert typing issues in split coequalizers has a similar, purely category theoretic interpretation. Split coequalizers are absolute in the sense that they are preserved under any functor. Such structures can be thought of as "diagrammatic," in the sense that their universal properties are expressed entirely in the language of category theory.
    ${ }^{15}$ See Theorem 2.31.

[^96]:    ${ }^{16} \mathrm{We}$ express these pointwise, where $F C \sim \gamma_{0}^{-1}(c)$.

[^97]:    ${ }^{17}$ Before moving on, it should be noted that, in the general case of an SPE, all the results in this section hold.

[^98]:    ${ }^{18}$ If the functor along which we are forming the relative comonad permits (right) Kan extensions, one obtains a form of composition [1]. But NF actually does a bit better.
    ${ }^{19}$ For the remainder of the proof, we adopt the category theoretic notation of modified-dependent products and indexed families presented as maps in slice categories.

[^99]:    ${ }^{20}$ The insertion of a union operation, to ensure $\hat{k}$ is appropriately "typed," seems problematic from the perspective of category theory, as $\cup$ is not functorial. But, it simply witnesses the fact that $T$ is a full and faithful embedding (i.e. $k\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right) \in T B$ implies $\left.\exists b \in B .\{b\}=k\left(\left\langle g_{j},\left\{c^{\prime}\right\}\right\rangle\right)\right)$.
    ${ }^{21}$ Recall, $g_{j}=\left\langle i d_{C_{1}}, \pi_{2} \circ g(-\circ j)\right\rangle$.

[^100]:    ${ }^{22}$ Recall the convention: $\hat{k}(\langle g,\{c\}\rangle) \equiv\langle\hat{k}(g),\{c\}\rangle$.

[^101]:    ${ }^{23}$ See $2.37,2.38,2.40$

[^102]:    ${ }^{24} \mathrm{But}$, there is no obvious means of recovering the original algebra, uniquely, from the resulting relative coalgebra.

[^103]:    ${ }^{25}$ Of course, this is also true for relative coalgebras whose base object is in the image of $F$. In particular, it is true of the free relative coalgebras.

[^104]:    ${ }^{26}$ The base fibre (i.e. the fibre over the terminal object of $\mathcal{N}$ ) of the externalization, $\mathcal{E}\left(1, C_{0}\right)$, is isomorphic to $T C_{0}$.
    ${ }^{27}$ Properties involving "structure" (e.g. (co)limits of diagrams) are slightly more complicated, but also permit abstraction to the fibred case [66].
    ${ }^{28}$ Sometimes referred to as a fully faithful adjoint string [36.

[^105]:    ${ }^{29}$ We include the "Identity" axiom, indicated by Diagram 4.1, although it is technically redundant [8, Lemma 3.2].

[^106]:    ${ }^{30}$ For the same reason, we could not have formed a KZ-monad for the restriction of Fam to small categories, Fam : Cat $\rightarrow C A T$.
    ${ }^{31}$ Strictly speaking, we would have a locally fully faithful relative KZ-(pseudo)monad, implying a relative version of a relationship proven in [69].
    ${ }^{32}$ This is proven in [8, Theorem 5.2].

[^107]:    ${ }^{33}$ Note: For $\mathbb{C} \in \operatorname{cat}(\mathcal{N})$, we will often refer to $\tilde{e} \mathbb{C}$ as $F a m \mathbb{C}$, as the two coincide in NF.

[^108]:    ${ }^{34} \mathrm{As}(A)_{K} \in F a m \mathbb{A}$, we associate the corresponding map $K \rightarrow A_{0}$ with $T K \rightarrow T A_{0}$.
    ${ }^{35}$ And the fact, trivial for NF, that $\{*\}$ is strongly cantorian.

[^109]:    ${ }^{36}$ Note: $!_{A}:(A)_{1} \rightarrow(A)_{1}$ is just the identity map.

[^110]:    ${ }^{37}$ The classical (unstratified) case is due to Kock [25].

[^111]:    ${ }^{38}$ For a standard KZ-monad $D$, the equivalence between $D$-cocomplete objects and pseudoalgebras is recorded in 69, Proposition 6].

[^112]:    ${ }^{39}$ As mentioned above, the converse to Lemma 4.37 holds in the standard case, has no apparent analogue in the relative case (see Lemma 4.44).

[^113]:    ${ }^{40}$ The modification $\lambda_{A}: T y_{A} \Rightarrow y_{T A}$ is central to Kock's definition of a KZ-Doctrine. Using the definition we have adopted, and where $\eta: 1_{T^{2}} \Rightarrow y_{T} \circ m$ is the unit of the right adjunction in $T y \dashv m \dashv$ $y_{T}$, we can show:

    $$
    \lambda_{A}=\eta_{A} \circ T\left(y_{A}\right)
    $$

    $$
    { }^{41} \chi=a \circ T a \circ T m_{A} \circ \lambda_{T A} .
    $$

[^114]:    ${ }^{42} y$ is no longer normalized.

[^115]:    ${ }^{43}$ Recall, this is an identity in the normalized case.

[^116]:    ${ }^{44}$ Notice this is comparable to the result that, given a section (up to isomorphism $\sigma$ ) of any component of $y$, one of the two triangle identities for $\hat{\sigma}$ and $\sigma$ holds automatically.

[^117]:    ${ }^{45}$ This is basically the "free" version of condition (2) of Definition 4.35

[^118]:    ${ }^{46}$ We know there are cases where the result holds, namely in the case of a relative pseudomonad along an identity functor.
    ${ }^{47}$ Local smallness with respect to the codomain fibration would equate to local cartesian closure. In $\mathcal{N}$, we have only modified local cartesian closure. See Propositions 3.59 and 3.61 .
    ${ }^{48}$ For further background, see [67, Section 9.6].
    ${ }^{49}$ As stated in the Declaration, Theorem 4.48 appears in collaborative work by Forster, Vidrine and

[^119]:    the author [14, Theorem 27]. Vidrine, in particular, should be acknowledged for this result.
    ${ }^{50}$ This reduces to the familiar homset, $\mathcal{C}(A, B)$, in the case that $A$ and $B$ are singleton families.

[^120]:    ${ }^{51}$ The idea of an internal topos is originally due to Bènabou [2]. The word "modern" is a direct reference to [67, Section 9.6]. It is the latter development, in the context of dependent type theory, that serves as the basis for what we apply to NF, in the remainder of this section.
    ${ }^{52}$ If $\mathcal{C}$ has a terminal object, the former coincides with the latter, $X$ is just the type-in-context $!_{X}: X \rightarrow 1$.

[^121]:    ${ }^{53}$ The relationship between cartesian closure and "paradoxical structures" is developed in 43 and 30.

[^122]:    ${ }^{54}$ Strictly speaking, the universal type would be the object of objects, $\tilde{\mathcal{N}}_{0}=V$. The classifying type is the object of $\mathcal{N} / V, \Gamma: \in_{\mathcal{N}} \rightarrow V$, which generates the full internal subcategory $\tilde{\mathcal{N}}$. Thus, we speak of the latter as an internal "universe" in the same way one would speak of a Grothendieck Universe (a full internal topos 67]).

[^123]:    ${ }^{55}$ This connects to the Remark on Curry-Howard, in Section 3.1.3 regarding the interpretation of $T B$ as the type of "proofs of $B$."
    ${ }^{56}$ This distinction appears in algebraic set theory, where one requires the adjoint $\exists \dashv(-)^{*} \dashv \forall$ (i.e. restriction to subobjects), but not the existence of the adjoint defining local cartesian closure $\sum \dashv(-)^{*} \dashv \Pi$. The latter, given the existence of a universe object $U$, would imply a contradiction.

[^124]:    ${ }^{58}$ So, in some sense, the class of internal dependent types is a proper subclass of the external dependent types.
    ${ }^{59}$ We write " $\mathrm{Fam}_{T}$ " to denote families indexed by objects of $T \mathcal{N}$.

[^125]:    ${ }^{60}$ Part of the broader goal of this thesis (and earlier work in [14) is to determine an appropriate, more general categorical structure, of which $\mathcal{N}$ is a particular example. Thus, we are particularly interested in cases where a syntactically convenient/obvious isomorphism turns out to be induced by a morphism with categorical/semantic significance.

[^126]:    ${ }^{61}$ Cantor's Theorem proves $|\mathcal{N}(1, V)| \prec|V|$.

[^127]:    ${ }^{62}$ One does not generally require the sumset of a class-indexed family of sets to be a set.

[^128]:    ${ }^{1}$ Earlier, we developed a form of stratified evaluation: modified-cartesian closure. But this is not sufficient for a standard combinatory algebra - consider the resulting Curry-Howard correspondence.
    ${ }^{2}$ Forster's suggestion is more direct. NF provides a setting where $|V| \cong|V \Rightarrow V|$, so one might hope to obtain a calculus of all functions in NF. No model has yet suggested itself, and it is unlikely that one could be obtained using the methods we develop here. In particular, the complementation operator is anti-continuous.
    ${ }^{3}$ Shortly after the publication of [55, both Plotkin and Scott provided a broad class of models (of which the multi-relation model is a particular case) which can be implemented in ZF 44,58 .

[^129]:    ${ }^{4}$ See Section 5.4
    ${ }^{5}$ This is a standard construction: $\langle x, y, z\rangle \sim\langle x,\langle y, z\rangle\rangle$. For its implementation in NF, see [50]. For an extension to steams ( $\omega$-sequences), see [10.
    ${ }^{6}$ Quine pairs allows for the unique correspondence between sets and ordered pairs. In this way, NF is said to have "surjective pairs." Using Quine sequences, we obtain a unique correspondence between sets and $\omega$-sequences. In turn, we can form a quotient, whereby there is a unique correspondence between sets and countable sequences of varying (i.e. possibly finite) length.

[^130]:    ${ }^{7}$ In other words, there is redundancy in our coding method. Each set corresponds to a unique sequence, but certain sequences (e.g. pairs) are coded by multiple sets.
    ${ }^{8}$ The identification of a stream terminating in a repeating sequence of $\emptyset$ 's with a finite sequence is discussed in [59].

[^131]:    ${ }^{9}$ In [58], for example, the subset of all coded sets is simply Fin.
    ${ }^{10}$ Plotkin's introduction of "graph models" formed the first investigation into general, set theoretic models of $\lambda$-calculus [44. Our work in Section 5.5 is both more and less general than that of Plotkin. We de-emphasize the role of "structure" (e.g. continuous lattice structure) and focus more on the role of comprehension (which is to say we assume the coding operations are definable the language of set theory).
    ${ }^{11}$ See [55, Definition 1.1].

[^132]:    ${ }^{12}$ See [55, Theorem 2.12].
    ${ }^{13}$ The function space, $[D \rightarrow D]$, inherits the lattice structure from $D$ in the expected way.

[^133]:    ${ }^{14}$ Scott uses the notation ' $P \omega$.'
    ${ }^{15}$ ' $e_{n}$ ' denotes the finite set coded by the exponents of the binary expansion of $n$. Similar methods allow for the coding of ordered pairs.

[^134]:    ${ }^{16}$ One could argue that this is the most natural interpretation of NF sets as functions, as each set corresponds to a unique distributive function. See Lemma 5.21 ,
    ${ }^{17}$ Note: Where the meaning is clear, we may refer to $a p p(A, B)$ as $A(B)$.

[^135]:    ${ }^{18}$ In the context of potentially infinite coded sets, it is worth noting a further advantage of coding sets via sequences. The sequence(s) coding a given set are endowed with an ordering relation between their members that need not hold between the elements of the coded sets.
    ${ }^{19}$ See 10 .

[^136]:    ${ }^{20}$ Recall how the "set-building" aspect of comprehension provides canonical examples of sets defining certain universal properties.

[^137]:    ${ }^{21}$ An inhomogeneous function $V \rightarrow V$, which is just the curried form of a homogeneous function $V^{n} \rightarrow V$, is not considered "inhomogeneous" (see Example 5.18).

[^138]:    ${ }^{22}$ Stratification of the multi-relation requires implementation of homogeneous finite sequences.
    ${ }^{23} \in \mathcal{C}$ indicates that the univariate function has a standard continuous graph in NF.

[^139]:    ${ }^{24}$ Among other shortcomings: a multivariate function that is distributive "point-wise" (i.e. in each variable) need not be distributive in its entirety (57]; neither $S$ nor $K$ are distributive.
    ${ }^{25}$ It should not be lost on anyone, just how striking it is for a set theory to model any of its total functions, much less form a set into which one can inject the collection of all total functions.

[^140]:    ${ }^{26}$ The image of a total idempotent function is comprised of its fixed points.

[^141]:    ${ }^{27}$ Much like $A \Rightarrow B$ in $\mathcal{N}$, the function space combinator, $A \rightsquigarrow B$, need not be an exponential object in $\mathcal{R}$.

[^142]:    ${ }^{28}$ The relevant comparison is the universal closure operation, $\lambda a \cdot \lambda x \cdot Y(\lambda y \cdot x \cup a(y))$, where $Y$ denotes the (non-distributive) paradoxical combinator [57, Theorem 5.6].

[^143]:    ${ }^{29}$ In a sense, the equivalence uses information "inside" the objects, rather than working at the level of arrows between objects.

[^144]:    ${ }^{30}$ See Definition 5.53, in the context of streams.
    ${ }^{31}$ See Corollary 5.49 .
    ${ }^{32} \mathrm{We}$ can do this in a couple ways. We could form a distinct projection function $\pi_{n}:\{1, \ldots, n\} \times V \rightarrow$ $V$, for each $n \in N$ (i.e. each distinct length of sequence) and then define $\bar{\pi}(n,-,-)$ as $\pi_{n} \cup\{\langle m, \emptyset\rangle|m\rangle$ $n\}$. In fact, the same method is used in Scott's formation of sequences in $P N$ [57, Formulas 2.19-2.23]. Alternatively, following [10] and Section 5.4 we could obviate any need for distinct projection function and interpret each set as a stream. $\pi: N \times V \rightarrow V$ can then be defined recursively (see Definition 5.53).

[^145]:    ${ }^{33}$ We could have included the empty sequence in $I$ and $K$. But, even without it, $I(\emptyset)=\emptyset$ and $\forall A . K(\emptyset)(A)=\emptyset$.

[^146]:    ${ }^{34}$ ' $B_{\text {Fin }}$ ' denotes the finite Boffa operation $x \mapsto\{y \mid x \in y \wedge y \in$ Fin $\}$.

[^147]:    ${ }^{35}$ In the sense that the necessary combinators are derivable from the abstract "theory of multi-relation $\lambda$-calculi," defined in 58, and satisfied by any model of NF.

[^148]:    ${ }^{36}$ Recall the formal version of application and abstraction, using Definitions 5.35 and 5.36

[^149]:    ${ }^{37}$ The original context in which Scott defines the multi-relation corresponding to $S$ assumes a counting of the universe [58, Section 4].

[^150]:    ${ }^{38}$ It is worth noting, the same identity: $\left\langle x_{0}, \ldots, x_{n}, \emptyset\right\rangle \sim\left\langle x_{0}, \ldots, x_{n}\right\rangle$ is employed in 57, Formulas 2.19 $-2.23]$.

[^151]:    ${ }^{39}$ It is interesting to note that one of the pairing functions described in [57. Section 5] has precisely these properties.
    ${ }^{40}$ As such, we typically state the latter as a prerequisite.

[^152]:    ${ }^{41}$ The difference exists only up to the elements of members of a given sequence, which are contained in $N$. Thus, the complications we encounter appear related to "coding" elements not being disjoint from the objects they are intended to code.

[^153]:    ${ }^{42} \mathrm{~A}$ result we extended to successor ordinals, in Lemma 5.50

[^154]:    ${ }^{43}$ We use $\downarrow$ to clarify the syntax. Unlike $\tau, \downarrow$ is inhomogeneous, so it does not have a graph in NF.

[^155]:    ${ }^{44}$ In fact, no multi-relation model will satisfy such a condition.

[^156]:    ${ }^{45}$ This identification follows automatically from the definition of pairs/sequences advocated by Scott and McCarty [59, Definition 4.1]. It also appears in Scott's earlier work [57, Formulas 2.19-2.23].
    ${ }^{46} \mathrm{We}$ are writing $V^{\omega} / \sim_{\omega}$ to emphasize the interpretation, we do not mean it as a formal quotient set (i.e. no pairs of distinct streams are identified under under $\sim_{\omega}$ ).

[^157]:    ${ }^{47}$ A stream can still "code" a finite set so, for the purposes of a multi-relation model of some form of $\lambda$-calculus, the informality of the quotient is not issue. For example:

    $$
    \{1,2,4\} \sim\langle 4,2,1,1,1, \ldots\rangle
    $$

    ${ }^{48}$ We used $\chi$ to denote a similar coding operation in Section 5.3 . See Definitions 5.35 and 5.36 .

[^158]:    ${ }^{49}$ This seems similar to an observation we made, regarding universal properties with universal sets. It might be worth considering if this similarity is, in any sense, "formal."
    ${ }^{50} \mathrm{We}$ can extend $\omega$-sequences, using limit and successor operations. Furthermore, there is no reason we could not drop one of the containment requirements. We could conceive of a homogeneous mapping of a set $x$ to a set of codes of $P_{\kappa}{ }^{\prime} x$ (resp. $B_{\kappa}$ ), where $\kappa$ is some aleph, and $P$ denotes the $\kappa$-powerset (equally, we could define a $\kappa$-superset). Thus, our interest in this homogeneous coding of sets extends beyond combinatory algebra to the general study of stratified set theories, namely, NF.
    ${ }^{51}$ We can still abstract continuous functions in way that preserves $\beta$-equivalence for non-empty sets.

[^159]:    ${ }^{52}$ Earlier, we observed there was a sense in which the distributive calculus was "natural" for Quine pairs. There will be a sense in which a calculus of $\omega$-continuous functions is natural for Quine sequences.

[^160]:    ${ }^{56}$ While the author has not seen any structure directly resembling what we refer to as "coding trees," it seems likely that similar ideas have been developed elsewhere in the (possibly unpublished) literature. As such, the author claims no priority in the general use of tree structures to understand coding of multivariate functions.
    ${ }^{57}$ Again, we can compare these directly with Definitions $5.34,5.35$ and 5.36

[^161]:    ${ }^{58}$ We have seen this earlier, in the form of $\eta$-equivalence. In a more elementary sense, it is just the assertion: $\langle\alpha, \chi\rangle: \mathcal{A} \rightarrow \mathcal{A} \times \mathcal{A}$ is injective.
    ${ }^{59}$ Such a pair would interest us, for example, if we were looking to $\lambda$-abstract a function $f$ and $y \in f^{6} z$.

[^162]:    ${ }^{60}$ The relevant fibre is equivalent to the intersection of fibres: $\alpha^{-1}(y) \cap \chi^{-1}(z)$.
    ${ }^{61}$ In the context of Example 5.70, $\gamma(x)=\operatorname{set}^{-1}(x)$.

[^163]:    ${ }^{62}$ For multivariate functions, we need to go as many levels down as there are variables.

[^164]:    ${ }^{63}$ It should be noted, a generative tree only determines the existence of a root, it need not be unique.
    ${ }^{64}$ In some ways, we are making an implicit assumption, regarding comprehension in our set theory.
    We assume that we can form the set corresponding to the " $\chi$-continuous fragment" (see Definition 5.79). This is unproblematic in NF, although we need to pay attention to the degree of inhomogeneity in our coding operations.
    ${ }^{65}$ For NF, we are interested in the case where $\chi$ is type-raising and $\alpha$ is homogeneous.

[^165]:    ${ }^{66}$ This is equivalent to the condition $\forall x \cdot \gamma(\{x\}) \neq \emptyset$.

[^166]:    ${ }^{67}$ Of course, it is also the case that $a(u)(b(u))=\bigcup_{x \in a}\{x\}(u)(b(u))$.

[^167]:    ${ }^{68}$ Even though we have reduced $b$ and $u$ their "relevant" members, the assignment:

    $$
    \chi \alpha(w)=b \text { and } \chi \alpha^{2}(w)=\chi(s) \cup \bigcup_{y \in b} \chi(y)
    $$

    requires (unreasonably) strong assumptions. The property $\alpha$ " $b=\chi \alpha(s)$ (even where $b=b_{s}$ ) does little to restrict the size of $b$.

[^168]:    ${ }^{69}$ Multiplicative Axiom: For any set $X$ of pairwise disjoint (non-empty) sets, there exists a set $C$ such that the for each $x \in X, C \cap x$ is a singleton set [52].
    ${ }^{70} \mathrm{By} \chi$-indexed partition, we mean $\coprod_{x} b_{x}=b \wedge \exists z \cdot \chi(z)=x$.

[^169]:    ${ }^{71}$ Any pre-combinatory algebra satisfying the conditions of Theorem 5.91 (specifically, Conditions 1 and 3) has a coding at least as strong as Scott's coding of finite sets 57.

[^170]:    ${ }^{72}$ If we do not wish to assume distributivity, we could consider a slightly more general version of Definition 5.92 where we do not assume $v \cap b_{z}$ is a singleton set, but rather a "coded" set (i.e. a member of $\left.P_{\chi}\left(b_{z}\right)\right)$. Assuming this principle:

[^171]:    ${ }^{73}$ See Remark after Theorem 5.44
    ${ }^{74}$ While we have expressed a preference for closure properties, such as Definition 5.90, as proxies for "smallness," we refer to "size" in the sense of cardinality.
    ${ }^{75}$ The countable multi-relation model in NF is one such example. A more general class of models could be obtained for any ordinal $\kappa$, permitting permitting a homogeneous coding of sets with cardinality $\preceq|\kappa|$.

[^172]:    ${ }^{76}$ An alternative approach to Section 5.5 could be to carry out Scott's model in the context of Maurer's Universes in Topoi 39. A natural numbers object satisfies the closure conditions of a universe object, as defined by Maurer. Furthermore, by exchanging "finite" ([57) for "small" (39]), one can develop much of the machinery developed by Scott, for $P(N)$ (the relationship between "finite" and "small" can be stated more formally: $\omega$ is an initial algebra of $\left.P_{\aleph_{0}}\right)$.

