SHEAF INTERPRETATIONS FOR GENERALISED PREDICATIVE INTUITIONISTIC SYSTEMS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

By
Nicola Gambino
Department of Computer Science

Contents

List of Tables

Abstract

Over the past century, formal systems based on intuitionistic logic have been the focus of much research in mathematical logic; in recent years, however, they have attracted interest also in computing because of their applications to computer-assisted formalisation of mathematics. The class of intuitionistic formal systems that are **generalised predicative** has emerged as particularly relevant both in mathematical logic and in computer science, as witnessed by the interest in Myhill-Aczel constructive set theories and Martin-Löf pure type theories.

A wide gap in our knowledge of this class of systems is due both to a scarce development of **sheaf interpretations** at the generalised predicative level, and to an unsystematic account of the connections between type theory, set theory and category theory. This gap prevents us from obtaining proof-theoretic results and from gaining conceptual insight into generalised predicative systems. The present work sets out to improve on this unsatisfactory situation.

The first main result of this thesis is the definition of sheaf interpretations for constructive set theories. We consider two kinds of sheaf interpretations: the first one is more explicit and lends itself more directly to applications, while the second one is more abstract and offers a clearer conceptual picture. We also apply sheaf interpretations to obtain relative consistency and independence proofs for constructive set theories.

A generalised type-theoretic interpretation of constructive set theories constitutes our second main result. To obtain it, we introduce logic-enriched type theories, which extend Martin-Löf pure type theories with judgements to express logic. One of the reasons for the interest in this generalised interpretation is that it leads to a precise match between set theories and type theories. Finally, we show how logic-enriched type theories can accommodate reinterpretations of logic, as inspired by sheaf interpretations in set theory.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

Copyright

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the head of Department of Computer Science.

Acknowledgements

First and foremost, I would like to thank my supervisor, Peter Aczel. I first benefited from his scholarly teaching in the academic year 1997/98, when he visited the University of Padua. I then enjoyed his stimulating guidance at the University of Manchester both in the academic year 1998/99, when I was a visiting student, and over the last three years, during which I worked on the research described here. To him, I wish to express my most sincere gratitude.

I acknowldedge financial support from the Department of Computer Science of the University of Manchester, the Mittag-Leffler Institut of the Swedish Royal Academy of Sciences and the Fondazione Aldo Gini of the University of Padua. In particular, I am very grateful to the organisers of the Mathematical Logic Year at the Mittag-Leffler Institut for their invitation.

I would like to thank David Rydheard and Andrea Schalk for their valuable advice, and especially Harold Simmons for his expert explanations on pointfree topology. I wish to express my gratitude also to Giovanni Sambin and Silvio Valentini for many inspiring lectures and interesting discussions on formal topology. I am thankful to Steve Awodey, Andrej Bauer, Venanzio Capretta, Thierry Coquand, Laura Crosilla, Giovanni Curi, Martin Escardó, Martin Hyland, Maria Emilia Maietti, Per Martin-Löf, Ieke Moerdijk, Erik Palmgren, Michael Rathjen, Andre Scedrov, Peter Schuster, Helmut Schwichtenberg, Dana Scott and Alex Simpson for stimulating suggestions and valuable comments on different aspects of the research described here. Robin Adams, Don McInnes and Ivar Rummelhoff were helpful in the final stages of the preparation of this thesis.

My father, my mother and my brother Michele constantly supported me and patiently put up with the physical distance that divided us in the recent years. I would like to thank also all my friends in Padua, Manchester, Edinburgh and Stockholm. Among them, I would like to mention especially Daniele for his visits. Last but not least, I thank Cristina, whose encouragement I cherish the most.

Chapter 1

Introduction

1.1 The variety of intuitionistic systems

Intuitionistic mathematics is not a uniform body of knowledge. A number of different, often competing, styles of intuitionistic mathematics have stemmed from Brouwer's pioneering work [13]. The plethora of alternative approaches is reflected in the variety of systems that have been formulated to support the development of intuitionistic mathematics [10]. The present work sets out to investigate intuitionistic formal systems. The motivations for our undertaking derive mainly from two areas: mathematical logic and theoretical computer science. Research concerning intuitionistic systems has led to a deeper insight into the differences between alternative approaches to intuitionistic mathematics and has been one of the central themes of mathematical logic over the past century. More recently, however, intuitionistic systems have also become of interest in theoretical computer science, where they have been applied to the formalisation of mathematics and to the verification of software.

The **propositions-as-types** idea has played a crucial role in gaining insight into the informal explanations for the axioms of intuitionistic logic [44, 76] and may indeed be considered to represent the essence of the 'constructive' content that intuitionistic proofs claim to have. We therefore reserve the word **constructive** for formal systems in which the propositions-as-types idea is assumed. We instead use the notion of **intuitionistic** system to mean something more general, where other treatments of logic are also possible. A key difference between constructive and intuitionistic systems is that various choice principles are allowed in the former, but not in the latter. Choice principles are indeed justified by the propositions-as-types idea [59], but not in general. For example Martin-Löf pure type theories are constructive systems [65], while the internal logic of an elementary topos is simply an intuitionistic system [12]. This thesis will focus mainly on intuitionistic systems: one of our aims is indeed to indicate that it is fruitful to allow other treatments of logic apart from the propositions-as-types one.

The systems we will consider are part of a family that is of particular relevance for intuitionistic mathematics. In order to isolate this family, it is convenient to consider an extension of the Weyl-Feferman-Schütte notion of **predicative** system [26]. Recall, for example, that the pure type theory $\mathbf{ML}_{<\omega}$ is a predicative system, while its extensions with rules for W-types, i.e. types of well-founded trees, are not [40]. We adopt the notion of **generalised predicative** system to mean systems that allow generalised inductive definitions and generalised reflection. For instance, extensions of pure type theories with W-types, as discussed above, and Mahlo universe types are generalised predicative systems [67]. Remarkably, the exact upper bound for the notion of generalised predicative system is however still to be determined. We will therefore introduce a working definition and say that a system is **fully impredicative** if the system of second-order arithmetic is proof-theoretically reducible to it [27]. For example, extensions of Martin-Löf pure type theories with the rules for a type of propositions in the style of the Calculus of Constructions [18] and the internal logic of a topos with a natural number object are fully impredicative systems [56].

In the following, we will investigate generalised predicative systems. Considerable evidence suggests indeed that large parts of classical and intuitionistic mathematics may be developed at the generalised predicative level [28]. In spite of this, there is a wide gap in our knowledge of generalised predicative systems. For example, many of the techniques available in the study of fully impredicative systems, such as sheaf and realisability interpretations, have not been developed yet for generalised predicative systems. As a consequence of this situation, many problems concerning these systems are still open. For example, important questions on the status of real numbers in intuitionistic mathematics are still to be settled [25]. Addressing these theoretical issues might contribute also to concrete applications. As already happened in the past [42, 55], a better understanding of the treatment of logic in intuitionistic systems could lead to further developments in the formalisation of mathematics using proof-checkers.

One of the challenges in the study of intuitionistic formal systems is that they may arise in three possible settings: type theory, set theory and category theory. There are good reasons to consider all of these settings, and the interplay between them will play a crucial role in this thesis. Type theory allows to express most directly foundational ideas and is best suited for the implementation in computer systems. Set theory, instead, is closest to idealised mathematical practice and provides a convenient setting for developing informally intuitionistic mathematics. Finally, category theory is most abstract and allows us to capture some of the essential properties of intuitionistic formal systems in a mathematically efficient way.

1.2 Sheaf and realisability interpretations

When considering intuitionistic formal systems, it is essential to distinguish between **semantical explanations** and **interpretations**. By means of illustration, we sketch this distinction for first-order intuitionistic logic. A semantical explanation of intuitionistic logic consists in a collection of informal explanations that serve to introduce and justify its axioms. The Brouwer-Heyting-Kolmogorov discussion of intuitionistic connectives is an example of semantical explanation, since it is intended to justify the assumption of the axioms of intuitionistic logic and explain the rejection of the law of the excluded middle.

By an interpretation we mean instead a formal assignment of mathematical objects to propositions, together with a notion of validity that specifies what it means for an axiom to hold under the interpretation. The well-known assignment of open sets of a topological space to propositions is an example of interpretation. It goes without saying that the notions of semantical explanation and of interpretation can be extended to formal systems that are more complex than first-order logic. For example, Martin-Löf's meaning explanations for his pure type theories are semantical explanations that are intended to convince us of their consistency [59]. Similarly, we can extend the notion of interpretation and, for example, speak of interpretations of type theories.

It is straightforward, but important, to observe that when we define an interpretation, we are developing some mathematics. For example, the topological interpretation of intuitionistic logic is usually defined working in set theory, manipulating open sets and operations on them. The question of which mathematical principles are used to define interpretations is rather subtle, since interpretations are often defined working informally in an unspecified formal system. If the formal system in which we are working informally coincides with the one whose interpretation we are defining, then we will speak of a **reinterpretation**. Kleene's realisability interpretation of Heyting arithmetic is an example of reinterpretation [52].

Of course, the definition of the interpretation cannot be fully formalised within the theory that is being interpreted because of Tarski's theorem on non-definability of truth and Gödel's incompleteness theorems [53, Chapter I, §14]. Considerations analogous to the ones that apply to Boolean-valued interpretations of classical set theories [11, pages 14 – 15] apply however to intuitionistic formal systems and reinterpretations may therefore be considered as leading either to syntactic translations or model-theoretic constructions.

Reinterpretations are an essential tool in the study of intuitionistic formal systems. They provide mathematically precise methods to establish proof-theoretic properties of formal systems, such as relative consistency and independence results. There are also conceptual reasons for the interest in reinterpretations, since they can be used to corroborate the informal semantical explanations given to introduce a formal system. Using a similitude, one might also say that reinterpretations show how formal systems behave like other mathematical structures in that they are 'invariant by change of base' [17].

Two main classes of reinterpretations have been considered in relation to intuitionistic formal systems. These are distinguished according to the mathematical structure that underlies their definition. On the one hand, we have **realisability interpretations**, i.e. interpretations whose underlying structure is a partial combinatory algebra or variations thereof [82]. On the other hand, we use the notion of **sheaf interpretations** to include interpretations based on topological spaces, complete Heyting algebras and sites, i.e categories equipped with a coverage [49]. These two families of interpretations have been considered in all of the three settings of type theory, set theory or category theory, but with different depth and alternative methods. Syntactic methods have been generally used in type theory and set theory, while categorical methods have been used in category theory.

Both sheaf and realisability interpretations seem to have been widely investigated for fully impredicative systems, while they have been rarely and unsystematically researched for generalised predicative ones. The absence of a comprehensive account of these interpretations at the generalised predicative level represents a serious obstacle to investigations into the proof-theoretic properties of generalised predicative systems, and prevents us from gaining further insight into the informal explanations that have been introduced to justify their introduction. This thesis aims at improving on this situation.

Let us conclude these general remarks on realisability and sheaf interpretations by pointing out that, although it is generally believed that only the former are related to the Brouwer-Heyting-Kolmogorov explanations for intuitionistic logic, also the latter bear some connection to the informal notion of 'proof' used in these explanations. We invite the interested reader to refer to [75] for further information.

1.3 On relating types, sets and categories

Until now, we have discussed interpretations within a fixed setting, but we may also consider the possibility of developing interpretations from one setting to another. At a first glance, the differences between type theory, set theory and category theory may seem to prevent this possibility. For example, note that while both type theories and the internal languages of categories may possess a rich type structure, set theories are generally formulated in a language with a unique and implicit type. Relating types, sets and categories is however possible and constitutes indeed one of the most important methods to extract information about intuitionistic formal systems. This is not an easy task: many of the intuitive, informal connections between the three settings do not seem to lead to precise statements, and the existing research on relating type theory, set theory and category theory does not seem to be completely satisfactory, even if it has been one of the main themes of research in intuitionistic systems over the past twenty-five years. An informal overview of these connections illuminates further motivations and goals of the work described in this thesis.

Deeply illuminating in its conceptual simplicity, the type-theoretic interpretation of Constructive Set Theory (CST) constitutes a fundamental link between set theories and type theories [2, 3, 4]. This interpretation presents however two aspects that motivate further research. Firstly, it rests on the assumption of the propositions-as-types treatment of logic in type theory. Secondly, when combined with a types-as-sets interpretation in the reverse direction, it leads to a mismatch between set theories and type theories [5]. These two aspects are actually related, and one might expect to obtain a precise match once the propositions-as-types treatment of logic is avoided.

The recent development of Algebraic Set Theory (AST) represents a robust approach to obtain connections between category theory and set theory [51]. Rather than attempting to isolate the categorical properties of categories of sets, as in Topos Theory, AST focuses on the categorical properties of categories of classes. The properties of sets are then described by axiomatising a distinguished class of maps, called small maps. While the axioms for the ambient category are meant to be fixed, axioms for small maps can be specified according to the set theory that is under consideration. At the fully impredicative level the connection between AST and Intuitionistic Set Theory (IST) is rather satisfactory, since we have axioms for small maps that correspond to intuitionistic set theories [51, 81] and that relate them with elementary toposes [8]. A first connection between set theories and categories at the generalised predicative level has been obtained in [62]. Although axioms for the ambient category and for small maps correspond well to the properties of the category of setoids in Martin-Löf pure type theories, the relationship with CST seems amenable of some improvements. A point originally remarked in [81] in relation to IST applies to CST as well. Although concrete categories of classes do not seem to be exact, exactness is one of the properties of the family of categories axiomatised in [62].

The relationship between type theory and categorical structures is a paradigmatic example of the difficulties in refining intuitive connections to establish precise statements relating different settings. The original connection between locally Cartesian closed categories and type theories [79] was very intuitive, and had the appeal of relating type theories to a well-known family of categories. Unfortunately, such a connection was flawed by subtle problems concerning the interpretation of substitutions in type theory. This problem was solved by replacing locally Cartesian categories with another family of categories [41, 43]. This family, however, does not seem to support in a natural way the categorical constructions that are available for locally Cartesian categories. This prevent us from transferring directly techniques from the categorical to the type-theoretic setting. Two approaches, motivated by different issues, have been proposed to recover the connection between well-known families of categorical structures and type theories. A first possibility is to shift the focus from categories of contexts to categories of setoids [42, 61, 62], while a second possibility is to formulate extensional type theories that are directly related to categorical structures [57, 58]. Let us finally mention that a potentially promising, but not yet widely explored, way to relate non-extensional type theories to categorical structures is suggested by the theory of exact completions [14, 15, 16, 60].

1.4 Aims of the thesis

The absence of a systematic account of sheaf and realisability interpretations and the problems in relating type theory, set theory, and category theory at the generalised predicative level motivates us to pursue a long-term project whose goals are:

(i) to isolate appropriate formal systems in the three settings,

(ii) to relate these formal systems with mutual interpretations,

(iii) to develop sheaf and realisability reinterpretations in each setting and to obtain proof-theoretic results as an application thereof,

(iv) to relate the reinterpretations developed in different settings.

The present thesis aims to contribute to this project. We will focus on sheaf interpretations for generalised predicative intuitionistic systems. Before describing our goals in more detail, we isolate the formal systems that we consider in the following and justify the reasons for the interest in them.

In set theory, our focus will be on the subject of Constructive Set Theory (CST). The pioneering works on CST were intended to formulate predicative or generalised predicative formal systems that could support the development of intuitionistic mathematics in set theory [35, 63]. More specifically, we will study set theories related to the formal system **CZF**, i.e. Constructive Zermelo-Frankel set theory [7]. The reasons for this choice are both mathematical and conceptual. First, the axioms for this set theory seem to correspond naturally to the principles required when developing generalised predicative mathematics in set theory. For example, the Subset Collection axiom, which is one of the peculiar axioms of **CZF**, seems to be necessary in order to show that the class of Dedekind cuts, as defined in CST, forms a set. A second reason to focus on systems related to **CZF** is the type-theoretic interpretation in Martin-Löf's pure type theories [2, 3, 4]. This interpretation allows us indeed to link notions and concepts from set theory to type theory.

In type theory, we will be interested in the area of Dependent Type Theory (DTT). The situation in this area is less straightforward than in set theory. On the one hand, there are Martin-Löf pure type theories [65]. Although these systems are generalised predicative systems, they treat logic using the propositions-as-types idea. On the other hand, there are the Calculus of Constructions and its extensions [18, 55]. Although these type theories do not assume the propositions-as-types treatment of logic, they are fully impredicative. In order to obtain generalised predicative intuitionistic systems, we will therefore be led to introduce type theories that extend Martin-Löf pure type theories so as to allow treatments of logic that are not informed by the propositions-as-types idea. We will call these formal systems **logic-enriched type theories**.

In category theory, it would be natural to explore Algebraic Set Theory (AST) [51]. This is because AST seems to be flexible enough to pursue the study of generalised predicative formal systems in a categorical setting, where recent research is currently aiming to isolate a notion of **predicative topos** [61, 62]. In this thesis the direct connections with AST will be rare, since we focus on the settings of type theory and set theory. Ideas and concepts of AST had however a remarkable influence on our work.

The first main goal of the thesis is the development of sheaf reinterpretations for CST. We aim to develop two kinds of sheaf interpretations: the first one will be based on complete Heyting algebras and the second one on sites. We also set out to obtain first examples and applications of sheaf reinterpretations, still working informally within constructive set theories. In order to do so, we develop some formal topology [68] in CST and reobtain some results of pointfree topology [30, 46] at the generalised predicative intuitionistic level.

Another main goal of the thesis is to indicate that a primitive treatment of logic in DTT is fruitful. Firstly, we will show that the type-theoretic interpretation of CST into pure type theories can be generalised to an interpretation into logic-enriched type theories. The main reason for the interest in the generalised interpretation is that it leads to a precise match when combined with a types-as-classes interpretation in the reverse direction. Secondly, we aim to show that logic-enriched type theories allow us to develop reinterpretations of logic that do not seem available for pure type theories. In particular, we will make a first step towards the definition of sheaf reinterpretations in DTT.

1.5 Methods

The proposed research is adventurous in nature. Let us highlight some of the challenges that we face, and discuss some of the ideas that we will exploit to address them.

1.5.1 Sheaf interpretations for CST

The main challenge to obtain sheaf interpretations for generalised predicative systems in the setting of set theory is to transfer results from Intuitionistic Set Theory (IST) to Constructive Set Theory (CST). Since Scott's topological interpretation of intuitionistic analysis [75, 77], the research on sheaf interpretations has mainly focused on the fully impredicative level, and hence on IST. While in IST the set existence axioms of classical set theories are retained and the other axioms are modified simply to accommodate intuitionistic logic [74], in CST some of the existence axioms of classical set theories are substantially modified or altogether dropped [7]. This is necessary in order to obtain generalised predicative systems instead of fully impredicative ones. For example, the Full Separation and the Power Set axiom, that are assumed in IST, are respectively modified and rejected in CST.

Although most working mathematicians might believe the Power Set and the Full Separation axioms are essential, this is not the case. Three ingredients allow to compensate their absence in CST: the class notation, the assumption of collection axioms, and the possibility of exploiting inductive definitions. Class notation will be used extensively when we work informally in CST. The assumption of the Power Set axiom, i.e. that the class of subsets of a set forms a set, can be avoided very frequently, provided that the formulation of some notions is reconsidered. For example, the carrier of a complete Heyting algebra is generally assumed to be a class in CST. The absence of Full Separation is instead partially compensated by the strengthening of the Collection axiom of IST to the Strong Collection axiom of CST [35]. The other collection axiom of CST that we will consider here is Subset Collection [7, Chapter 7], which allows us to recover some of the consequences of the Power Set axiom. For example, the class of Dedekind cuts can be proved to be a set under the assumption of Subset Collection [7, Section 3.6]. Finally, inductive definitions provide a very efficient and widely applicable method to replace the classical use of ordinals. For example, we will often isolate classes whose elements will be used to interpret the sets of CST using inductive definitions. Furthermore, assuming the Regular Extension axiom, it is also possible to show that a wide family of inductive definitions determine sets [7, Chapter 5].

When it comes to the definition of sheaf interpretations, the differences between IST and CST present us with two different issues. On the one hand, we need to see whether these interpretations can be actually be defined working informally in CST. On the other hand, we have to investigate whether the axioms of CST are valid under the interpretations. The first issue becomes clear once we recall that sheaf interpretations are generally developed using notions, such as that of a complete Heyting algebra, that need to be reconsidered in CST. The second issue is even more problematic, and is best illustrated considering sheaf interpretations of collection axioms of IST and CST. At the fully impredicative level, the validity of the Collection axiom of IST under sheaf interpretation is generally proved using the Full Separation axiom that is not assumed in CST. For example, this is the case in the double-negation translation of classical Zermelo-Frankel set theory into its intuitionistic counterpart [34]. Therefore, it is not clear a priori whether we can prove the validity of the Strong Collection axiom of CST without the assumption of Full Separation.

A careful distinction between classes and sets, on the one hand, and between arbitrary and restricted formulas, on the other hand, represents a promising approach to address the first and the second issue, respectively. We already mentioned the example of the notion of a complete Heyting algebra, saying that in CST we may allow the carrier of the algebra to be a class. One of the reasons to do this is to define a number of examples while still working in CST. Once the basic notions are set up, the development of pointfree and formal topology [46, 68] provide us indeed with guidance to obtain examples. When it comes to the definition of sheaf interpretation of formulas, arbitrary formulas will be associated to classes, and restricted formulas to sets. This is motivated by the simple, but important, observation that sentences of the language of CST correspond to subclasses of the class $\{\emptyset\}$, while restricted sentences correspond to subsets of it. It seems therefore sensible to respect this distinction when it comes to defining sheaf interpretations. As we will see in Part II of the thesis, the development of sheaf interpretations for CST exploits in an essential way not only these ideas, but also the assumption of the collection axioms in CST.

1.5.2 Collection principles in DTT

The propositions-as-types idea seems to be one of the main obstructions to the development of sheaf interpretations in DTT. On the one hand, considerable research in categorical logic shows that strong choice principles are not generally valid under sheaf interpretations [56, 62] and on the other hand a type-theoretic version of the axiom of choice is derivable assuming the propositions-as-types paradigm [59, pages 50 – 52]. To approach the study of sheaf interpretations in DTT, it seems therefore convenient to abandon the propositions-as-types idea. We will therefore formulate logic-enriched type theories, that are extensions of pure type theories in which logic is treated as primitive. The development of sheaf interpretations in DTT may then be obtained following the ideas that inspired the sheaf interpretations in CST, provided that we relate constructive set theories and logic-enriched type theories. We are therefore naturally led to consider a generalisation of the original type-theoretic interpretation of CST [2, 3, 4].

The generalisation, that involves replacing pure type theories with logic-enriched type theories, is not straightforward. The propositions-as-types idea plays indeed a crucial role to prove the validity of Restricted Separation, Strong Collection and Subset Collection axioms of CST under the original type-theoretic interpretation. In particular, the type-theoretic axiom of choice, that is no longer derivable in logic-enriched type theories, was exploited in the proofs regarding the collection axioms. To overcome its absence, we consider collection principles in DTT that are related to the collection axioms of CST.

The formulation of these collection principles is constrained by the non-extensional aspects of DTT [42, 64]. We will then have to resort to the expressive power allowed by type dependency and to the variety of constructs available in logic-enriched type theories. The non-extensional aspects of DTT reveal themselves to be a major obstacle also in the development of sheaf interpretations. We will therefore limit ourselves to considering only one of the two steps in which sheaf interpretations can be divided. We leave investigations into the other step for future research. Remarkably, the step that we consider can be accommodated also in the non-extensional setting of DTT. At this point, the development of sheaf interpretations for CST and the generalised type theoretic interpretation can be put into use. Once again, however, we have to overcome the non-extensional aspects of DTT.

1.6 Overview

This thesis is organized in three parts. Part I is mainly devoted to introduce constructive set theories and pure type theories, and to fix notation and terminology. Part II presents sheaf interpretations for CST. Part III introduces logic-enriched type theories and applies collection principles in DTT.

Part I includes chapters 2 and 3. Constructive set theories are introduced in Chapter 2, where we relate them with classical and intuitionistic set theories. The axioms for all the set theories considered in this thesis are contained in Appendix A. We also discuss some of the peculiar aspects of the development of mathematics in CST, and prove some of the results that will be used to develop sheaf interpretations for CST. Chapter 3 is just a review of the formulation of pure type theories, and is accompanied by Appendix B, where we spell out the rules for the type theories considered here.

Part II includes chapters 4, 5 and 6. Chapter 4 develops some formal topology in CST. First we isolate the notion of a set-generated frame and the notion of a set-presented frame, that are most relevant for this development. We then study concrete examples of these notions. This will be convenient in order to obtain examples and applications of sheaf interpretations. Chapter 5, based on joint work with Peter Aczel, studies Heyting-valued interpretations for CST [37]. At the end of the chapter, we present some first proof-theoretic applications and discuss further potential directions of research. Chapter 6 concludes Part II by presenting sheaf interpretations for constructive set theories as determined by a site. While developing these interpretations, we will often remind the reader of the analogy between the category of classes and the category of presheaves, that we actually make precise using some of the ideas of AST.

The results obtained up to this point will highlight that the the collection axioms of CST have a peculiar relationship with sheaf interpretations. On the one hand they play a crucial role in setting up these interpretations, and on the other hand they are preserved by these interpretations, i.e. they imply the validity of their instances under the interpretations. The introduction of logic-enriched type theories allows us to transfer some of these facts to DTT.

Part III includes chapters 7, 8 and 9 and it is based on joint work with Peter Aczel [6]. Only results due to the author of the present thesis are included here, unless it is necessary to do otherwise for reasons of clarity, and this will be indicated explicitly. Chapter 7 is devoted to introduce logic-enriched type theories as extensions of pure type theories and to formulate collection principles. Collection principles find their first application in Chapter 8, where we present a generalised type-theoretic interpretation of CST in logic-enriched type theories. Here, collection principles are used to prove the validity of the collection axioms under the interpretation. Chapter 9 studies how logic-enriched type theories accommodate reinterpretations of logic. We first focus on a propositions-as-types interpretation, that reduces a logic-enriched type theory to its pure counterpart. Collection principles are valid under the propositions-as-types interpretation, since they are consequences of the type-theoretic axiom of choice that is derivable in pure type theories. We then study reinterpretations of logic, as in one of the steps of sheaf interpretations. Here collection principles play a double role: on the one hand they are used to set up the reinterpretation and on the other hand they are preserved by it, just as in CST. We end the thesis with conclusions and a discussion of future research in Chapter 10.

Part I

Generalised predicative systems

Chapter 2

Constructive set theories

2.1 Language and axioms

In this chapter we introduce Constructive Set Theory (CST) and present some first results that will be used to develop sheaf interpretations. One of the particularly good features of constructive set theories is that they can be introduced very directly, like classical ones [53]. Furthermore, the development of mathematics in CST can take advantage of the class notation that is widely used in mathematical practice. In this section, we will set up a set-theoretic language and then present constructive set theories by simply listing their axioms. Since the axioms for a constructive set theory are very different from the ones for a classical or intuitionistic set theory, we will help the readers to gain some insight into CST by reviewing some of the connections between classical, intuitionistic and constructive set theories in Section 2.2.

We will then focus just on CST. Firstly, we fix some notation that will be used when working informally in CST and recall the familiar class notation that will be used extensively elsewhere in the thesis. Secondly, we will take a category-theoretic perspective and describe some of the properties of the category of classes. This will be useful when, in Chapter 6, we relate the category of classes to categories of presheaves. Finally, at the end of the chapter, we prove some results that will be particularly useful in the development of sheaf interpretations for CST. Overall, this chapter contains the minimal background needed to read the remainder of the thesis; we will often point the reader to specific sections and results of [7] that the reader might consult for more details.

Let us conclude these introductory observations with a remark concerning the use of the word 'constructive', that we decided to reserve for systems that adopt the propositions-as-types treatment of logic. Although logic is treated as primitive in set theory, the propositions-as-types idea is an essential component for the justification of the axioms of CST, as given by the original type-theoretic interpretation [2, 3, 4]. It seems therefore reasonable, as well as coherent with the terminology existing in the literature, to speak of constructive set theories.

Remark. Let us highlight that we use 'CST' as an abbreviation for 'Constructive Set Theory', i.e. the general area of research, while we use **CZF** for the specific formal systems of Constructive Zermelo-Fränkel set theory, that we isolate in what follows. We generally use abbreviations in boldface font to stand for specific formal systems.

2.1.1 Language

We now introduce the language $\mathcal{L}$ that will be used to formulate the axioms of the set theories we consider. We prefer to assume restricted quantifiers as primitive symbols and to have the membership relation as defined symbol. This will be convenient when we come to the definitions of interpretations of constructive set theories in Chapter 5 and Chapter 6.

The language $\mathcal{L}$ has symbols x_i for each natural number i, a binary relation symbol $=$ denoting equality, a constant $\bot$ denoting the canonical false proposition, binary connectives $\wedge$, $\vee$, $\rightarrow$ denoting conjunction, disjunction and implication respectively, unrestricted quantifiers $(\forall x_i)$ and $(\exists x_i)$, for all natural numbers i, denoting universal and existential quantification over all sets, and restricted quantifiers $(\forall x_i \in x_j)$ and $(\exists x_i \in x_j)$, for all natural numbers i and j denoting universal and existential quantification over sets that are elements of x_j. An **expression** is any list of symbols. We say that an expression is a **formula** if it is formed according to the following rules:

1. $x_i = x_j$ is a formula for all i and j,
2. $\bot$ is a formula,
3. if ϕ and ψ are formulas then $(\phi) \wedge (\psi)$, $(\phi) \vee (\psi)$ and $(\phi) \rightarrow (\psi)$ are formulas,
4. if ϕ is a formula, then $(\forall x_i \in x_j)\phi$ and $(\exists x_i \in x_j)\phi$ are formulas for all i and j ,
5. if ϕ is a formula then $(\forall x_i)\phi$ and $(\exists x_i)\phi$ are formulas for all i.

We say that a formula is **restricted** if it constructed using only 1 – 4 above. For x_i and x_j define the membership relation by letting

$$x_i \in x_j =_{\text{def}} (\exists x_k \in x_j) x_i = x_k \, .$$

Observe that the membership relation is defined by a restricted formula. We can then introduce some usual definitions: for ϕ, ψ formulas and x_i, x_j variables define

$$\begin{aligned} \top &=_{\text{def}} \bot \to \bot \,, \\ \phi \leftrightarrow \psi &=_{\text{def}} (\phi \to \psi) \wedge (\psi \to \phi) \,, \\ \neg\phi &=_{\text{def}} \phi \to \bot \,, \\ x_i \notin x_j &=_{\text{def}} \neg x_i \in x_j \,. \end{aligned}$$

We define $\mathcal{L}^{(V)}$ as the extension of the language $\mathcal{L}$ that is obtained by adding constants for sets. This language will be used in the following to state set-theoretic axioms.

Free and bound variables. Free and bound variables of formulas are defined as usual. We write FVϕ for the set of free variables of a formula ϕ.

We say that a formula ϕ is a **sentence** if it has no free variables.

Substitution. Let ϕ be a formula with $\mathrm{FV}\phi = \{x_1, \ldots, x_n\}$. We write

$$\phi[e_1, \ldots, e_n/x_1, \ldots, x_n]$$

for the result of simultaneously substituting expressions e_i for the free occurrences of x_i in ϕ for $i = 1, \ldots, n$.

When working informally in a set theory, it is convenient to depart from the strict definition of the formal languages $\mathcal{L}$ and $\mathcal{L}^{(V)}$ and use a richer notation. In the following we use lower-case Greek letters ϕ, ψ, ξ for arbitrary formulas, and θ, η for restricted formulas. We use symbols x, y, z, u, v, w to stand for the variables of $\mathcal{L}$. Finally, we use lower-case letters that are not used for variables to stand for constants for sets. As usual, we will make use of with accents, subscripts and superscripts when this is convenient.

2.1.2 Axioms for set theories

The axioms for a set theory can be naturally divided into two groups: logical axioms and set-theoretical axioms. Logical axioms specify the underlying logic of the set theory, e.g. classical or intuitionistic, while set-theoretical axioms specify properties of sets. As we will see in our discussion of set theories in Section 2.2 there is remarkable interaction between the two groups of axioms.

Let us briefly discuss logical and set-theoretical axioms that will be considered in the remainder of the chapter. We first review logical axioms. Constructive set theories have the usual axioms for first-order intuitionistic logic as well as axioms in (2.1) for restricted quantifiers. We write the logical axioms for restricted quantifiers as schemes in the extended language $\mathcal{L}^{(V)}$, and use the definitions and the notational conventions that we have just introduced:

$$\begin{array}{lcl} (\exists x \in a)\phi & \leftrightarrow & (\exists x)(x \in a \wedge \phi)\,, \\ (\forall x \in a)\phi & \leftrightarrow & (\forall x)(x \in a \rightarrow \phi)\,. \end{array} \tag{2.1}$$

The law of **excluded middle** (**EM**) is the scheme

$$\phi \vee \neg\phi\,,$$

where ϕ is an arbitrary formula. We may consider a variation over the law of excluded middle that is suggested by the distinction between arbitrary and restricted formulas. The law of **restricted excluded middle** (**REM**) is the scheme

$$\theta \vee \neg\theta\,,$$

where θ is a restricted formula. Let us now discuss set theoretical axioms. The set theoretic axioms can be naturally divided into three groups: structural, set existence and collection axioms. All the axioms for the set theories we consider are presented in Appendix A. They are formulated in the language $\mathcal{L}^{(V)}$. Observe that some of them are given as axiom schemes. When referring to an axiom we prefer to avoid to explicitly recall whether it is an axiom scheme or not, as this is clear from the name used to denote it. We will therefore simply say 'Pairing' or 'Strong Collection' instead of 'Pairing axiom' or 'Strong Collection axiom scheme'.

Constructive Zermelo-Frankel set theory (**CZF**) is the set theory with Extensionality and Set Induction as structural axioms; Pairing, Union, Infinity and Restricted Separation as set existence axioms; Strong Collection and Subset Collection as collection axioms. We also consider $\mathbf{CZF}^-$ and $\mathbf{CZF}^+$, that are a subtheory and an extension of **CZF**, respectively. Starting from **CZF**, the set theory $\mathbf{CZF}^-$ is obtained by omitting Subset Collection, while $\mathbf{CZF}^+$ is obtained by adding Regular Extension [7].

2.2 On relating set theories

Readers who are not familiar with CST may find difficult to understand its axioms. We will therefore briefly review some well-known facts relating classical, intuitionistic and constructive set theories, as they may help gaining some insight into the ideas motivating the formulation of the axioms of **CZF**. We assume the reader has some familiarity with set theory [53].

First of all, let us recall the axioms of classical and intuitionistic Zermelo-Frankel set theory. Classical Zermelo-Frankel set theory (**ZF**) is the set theory based on classical logic with Extensionality and Foundation as structural axioms; Pairing, Union, Infinity, Full Separation, and Power Set as set existence axioms; Replacement as the only collection axiom. Intuitionistic Zermelo-Frankel set theory (**IZF**) is the set theory based on intuitionistic logic with Extensionality and Set Induction as structural axioms; same set existence axioms of its classical counterpart; and Collection as the only collection axiom. Table 2.1 highlights the difference between the axioms of these three set theories.

Table 2.1: Axioms and set theories.

Axioms	**ZF**	**IZF**	**CZF**
Extensionality	✓	✓	✓
Foundation	✓		
Set Induction		✓	✓
Pairing, Union, Infinity	✓	✓	✓
Power Set	✓	✓	
Full Separation	✓	✓	
Restricted Separation			✓
Replacement	✓		
Collection		✓	
Strong Collection			✓
Subset Collection			✓

Observe that **CZF** is a subtheory of **IZF**, which is in turn a subtheory of **ZF**. **CZF** is included in **IZF** because Strong Collection follows from Collection and Full Separation, and because Subset Collection follows from Power Set [2]. **IZF** is included in **ZF** because, assuming classical logic, Set Induction follows from Foundation and Collection follows from Replacement [74].

2.2.1 Structural and set existence axioms

When considering structural axioms we contrast **ZF**, that assumes Foundation, with **IZF** and **CZF**, that assume Set Induction. There are two reasons for replacing Foundation with Set Induction in set theories based on intuitionistic logic. Firstly, assuming the law of excluded middle, Foundation is equivalent to Set Induction. Secondly, Set

Induction does not imply the law of excluded middle, while Foundation does [7, Chapter 9]. In the discussion of set existence axioms we contrast **CZF** on one side, with **ZF** and **IZF** on the other side. The reason for this is that **CZF** is a generalised predicative system, while both **ZF** and **IZF** are fully impredicative. In particular, the set existence axioms of **CZF** are obtained from the ones of **IZF** by dropping Power Set and replacing Full Separation with Restricted Separation.

Avoiding the use of Power Set is not as problematic as one might expect, since we can still treat the collection of subsets of a set as a class, as we will discuss in Section 2.3. Furthermore, some of the applications of Power Set in Classical Set Theory can be recovered by using Subset Collection that is part of **CZF**. The situation is similar with respect to Full Separation. We can of course treat collections of elements of a set determined by an unrestricted formula as classes, but it is a remarkable that when developing mathematics informally in set theory, Full Separation is rarely used in its 'full' form, and Restricted Separation often suffices, as we will see in Chapter 4. Furthermore, the formulation of Strong Collection partially compensates the absence of Full Separation from **CZF**, as we now discuss.

2.2.2 Collection axioms and REA

We now discuss the collection axioms of **CZF**. As we will see, these axioms play a crucial role in the development of sheaf interpretations for CST. Furthermore, they seem to provide much of the 'expressive power' that is used to develop constructive mathematics in set theory.

Let us first discuss Strong Collection. Observe that, using just the axioms of intuitionistic logic, Strong Collection implies both Collection and Replacement. Furthermore, assuming Full Separation, Collection implies Replacement. An analysis of the reverse implications provides some insight into the formulation of the Strong Collection axiom. In Zermelo set theory, i.e. **ZF** without Replacement, classical logic allows us to show that Replacement implies Collection, which in turn implies Strong Collection, using Full Separation. In Intuitionistic Zermelo set theory, i.e. **IZF** without Collection, Full Separation still allows us to prove that Collection implies Strong Collection. Without the assumption of classical logic, Replacement does not imply Collection, as proved by Harvey Friedman: we invite the reader to refer to [74] and references therein for more details. Within an intuitionistic set theory, one is therefore led to consider Collection as an axiom. Once Full Separation is omitted, however, Collection does not seem to imply Strong Collection anymore. To retain all the strength of the Replacement axiom scheme of **ZF**, it seems therefore that Strong Collection is the appropriate axiom to consider in CST. In what follows, when working informally in **CZF**, we will sometimes indicate explicitly applications of Strong Collection for which Replacement is sufficient.

Let us now discuss Subset Collection. First of all, recall that Power Set implies Subset Collection which in turn implies Myhill's Exponentiation axiom [63], expressing that the class of functions between two sets is a set. These implication are most easily seen by introducing the Fullness axiom that is equivalent to Subset Collection [7, Chapter 7]. Subset Collection has however mathematical consequences that do not seem to follow from Exponentiation, such as the sentence asserting that the Dedekind real numbers form a set [7, Section 3.6].

The Regular Extension axiom (REA) that is part of the set theory $\mathbf{CZF}^{+}$, will play an important role when we use inductive definitions in CST. Assuming REA, it is indeed possible to show that many inductive definitions determine sets [7, Chapter 5]. We will also make use of Set Compactness theorem that is a consequence of REA [7, Section 5.5].

2.2.3 Adding excluded middle

One of the important features of **CZF** is that if we extend it with the law of excluded middle (**EM**) we obtain a set theory that has the same theorems as **ZF**. This is because, assuming **EM**, Strong Collection implies Full Separation and Subset Collection implies Power Set. To prove this second claim, recall that Subset Collection implies Exponentiation, and observe that **EM** implies that the power set of $\{0\}$ is $\{0, 1\}$. Using these facts, for any set a we can form the set of the functions from a to $\{0, 1\}$. These functions behave like characteristic functions and thus give us the power set of a, as wanted.

We may also consider extensions of a constructive set theory with restricted excluded middle scheme (**REM**). A refinement of the argument just presented shows that if we extend **CZF** with **REM** we can still derive the power set axiom. We may observe that the classical set theory that is obtained from **ZF** by dropping replacement and substituting Full Separation with Restricted Separation is fully impredicative, because of the presence of Power Set. This set theory has a double-negation interpretation into its intuitionistic counterpart [34], which is a subtheory of $\mathbf{CZF} + \mathbf{REM}$. Therefore the extension of **CZF** with **REM** is a set theory that has at least the same proof theoretic strength as second-order arithmetic, and hence is fully impredicative [7, Chapter 9].

2.3 Working in CST

2.3.1 Classes

When developing constructive mathematics in CST we can take advantage of the familiar informal set theoretic notation that is used also in the classical context. This informal notation will be widely exploited in the remainder of the thesis. In particular, we will use classes and the convenient notation associated to them [7, Chapter 3]. We will use upper-case letters $A, B, C, \ldots$ to stand for classes. Here we list the definitions of some classes that will be used in the remainder of the thesis. Working informally within **CZF**$^-$, let us define the empty class and the class of all sets respectively as

$$\begin{aligned} \emptyset &=_{\text{def}} \{x \mid x \neq x\}\,, \\ \mathsf{V} &=_{\text{def}} \{x \mid x = x\}\,. \end{aligned}$$

The following definitions will be also useful

$$\begin{aligned} \mathsf{0} &=_{\text{def}} \emptyset\,, \\ \mathsf{1} &=_{\text{def}} \{\mathsf{0}\}\,, \\ \mathsf{2} &=_{\text{def}} \{\mathsf{0}, \mathsf{1}\}\,. \end{aligned}$$

Furthermore we use N and Q to stand for the classes of natural and rational numbers, respectively, that are defined in the usual way [7, Section 3.5]. We use $\mathsf{R_d}$ for the class of Dedekind reals, as defined in [7, Section 3.6]. In many important cases it is possible to show that a defined class is actually a set. The class of natural numbers, for example, can be shown to be a set using Infinity. For real numbers the issue is more complex, and therefore here we just recall that in **CZF** the class $\mathsf{R_d}$ of Dedekind cuts forms a set, and invite the reader to refer to Theorem 3.24 of [7] for more details.

2.3.2 Power classes

The distinction between sets and classes is a crucial one in CST, and it needs to be carefully treated when working informally within a constructive set theory. Let us point out two aspects of this distinction that will play an important role in the rest in the thesis: the distinction between subsets and subclasses and the use of power classes.

Let A and P be a classes. We say that a class P is a **subclass** of A if $P \subseteq A$, and that it is a **subset** if it is also a set. The distinction between subclasses and subsets is particularly relevant because the Full Separation and the Power Set are not part of the set theories we will work in. For a set a and a formula ϕ with a free variable x we may define

$$P =_{\text{def}} \{x \in a \mid \phi\}\,. \tag{2.2}$$

Observe that P is a subclass of a. Without the assumption of Full Separation it is not possible to assert that P is a set in general.

We may also consider the **power class** of a, i.e the class of subsets of a set a, defined as follows:

$$\mathsf{Pow}\, a =_{\mathrm{def}} \{x \mid x \subseteq a\} \,.$$

Without the assumption of Power Set, this class cannot be asserted to be a set. Observe that elements of $\mathsf{Pow}\, a$, just like elements of any other class, are sets. Hence, we cannot assert that the class defined in (2.2) is an element of $\mathsf{Pow}\, a$ unless we first show that it is a set.

2.3.3 Truth values

The power class of 1 will play an important role in the applications of Heyting-valued interpretations in Chapter 5 and therefore it is convenient to have a special notation for it:

$$\mathsf{P} =_{\mathrm{def}} \mathsf{Pow}\, 1 \,.$$

The distinction between classical, intuitionistic set theories can also be appreciated by contemplating the different status of the class P in the three contexts. For p in P, define

$$!\, p =_{\mathrm{def}} (\exists \_ \in p)\top \,,$$

where $\_$ stands for an anonymous bound variable. For for a sentence ϕ, define

$$[\![\phi]\!] =_{\mathrm{def}} \{\_ \in 1 \mid \phi\} \,,$$

and observe that we have

$$\phi \leftrightarrow\, ![\![\phi]\!] \,.$$

There are three separate aspects to consider when discussing the class P: first, whether P is a class or a set; second, what are the elements of P; and finally what algebraic structure P possesses. In **CZF** the class P cannot be asserted to be a set, and we must distinguish between subsets and subclasses of 1. Arbitrary and restricted sentences correspond to subclasses and subsets of 1, respectively and therefore, elements of P correspond only to restricted sentences. In **IZF**, Power Set allows us to prove that the class P is a set and Full Separation implies that we do not need to distinguish between subclasses and subsets of 1. Therefore elements of P correspond to arbitrary sentences. In **ZF** the situation is even more extreme and, by the law of excluded middle, we can show that P coincides with 2. The algebraic structure on P are also very different in the three contexts. We have that P is a complete Boolean algebra in **ZF** and a complete

Heyting algebra in **IZF**. As we will see in Chapter 4, P is instead an example of the notion of 'set-generated frame' within **CZF**.

2.4 Categories of classes

To gain further insight a particular part of mathematics, it is often helpful to take a step back from its concrete development, and stare at it from a more abstract perspective, such as the one given by category theory. This is not only conceptually important, but often mathematically efficient. It is however crucial to choose carefully which category one should focus on. For example, it has been argued that replacing the category of frames with the one of locales is a key step in pointfree topology [48]. Set theory is no exception in this respect. While categories of sets have been the focus of much attention, and provide one of the motivations for the introduction of Topos Theory [54], categories of classes have been largely ignored, at least until the recent development of AST, i.e. Algebraic Set Theory [51, 61, 62, 81].

One of the reasons to prefer considering categories of classes over categories of sets is their generality. Sets are indeed special kind of classes, namely classes that are 'small' in some sense. Secondly, there are concepts and axioms that can hardly be expressed considering only categories of sets. An important example of this phenomenon is the Collection axiom of Intuitionistic Set Theory [51]. Once the perspective of AST is assumed and categories of classes are taken as the basic object of study, however, we face the problem of isolating sets from classes. In line with other parts of category theory [50] or categorical logic [45], this can be done by axiomatising a distinguished family of maps, that in this case are called **small**. In concrete categories of classes, small maps are the functions between classes whose fibers are sets, as we will see in Subsection 2.4.2.

2.4.1 Categories of classes, abstractly

Here, we will only review some basic concepts of AST that will be useful in Chapter 6 to relate formally the category of classes with the category of presheaves, as defined in CST. We assume that the reader has some familiarity with categorical logic [66]. For more information on AST we invite the reader to refer to [51, 81] and in particular to [61, 62] for its aspects more closely related to generalised predicative theories. For the readers that are not familiar with AST, more intuition will be hopefully available after reading Subsection 2.4.2, where we describe a concrete example of the abstract notions described here.

A reasonable basic setting to consider axioms for small maps is the one of a category $\mathcal{E}$ that is regular, has stable disjoint coproducts, and has right adjoints to pullback functors between families of subobjects. Given a map F in $\mathcal{E}$, we write $\exists_F$ and $\forall_F$ for the left and right adjoint adjoints to the pullback functor F^* induced by F, where the existence of the left adjoint is a consequence of the fact that $\mathcal{E}$ is regular. We like to draw this situation as follows:

$$\begin{array}{ccc} A & \qquad & Sub A \\ \downarrow F & & \exists_F \left(\;\uparrow F^* \;\right) \forall_F \\ B & & Sub B\,. \end{array}$$

We can then assume to have a distinguished family $\mathcal{S}$ of small maps that satisfies, for example, the axioms for small maps listed in Section C.1 on page 194. We say that an object A is **small** if the map $A \longrightarrow 1$ is small, where 1 is a terminal object of the category $\mathcal{E}$.

The following notion will be of some interest both in categories of classes and in categories of presheaves in CST. Its formulation is taken from [51, Chapter I, §3], but it is also appears in [81], but under the name 'small relation'.

Definition 2.1. Let A be a an object of $\mathcal{E}$. For an object I of $\mathcal{E}$, we say that a subobject $R \rightarrowtail I \times A$ is an **I-indexed family of small subobjects of** A if the composite map $R \rightarrowtail I \times A \xrightarrow{\pi_1} I$ is a small map. ♢

This notion allows to describe in an abstract way the behaviour of power classes, as expressed by axiom (**P1**) of Section C.2. The next notion that we introduce is taken from [81] and will also be of interest both in categories of classes and in categories of presheaves.

Definition 2.2. We say that an object U in $\mathcal{E}$ is an **universal** if for every object A in $\mathcal{E}$ there is a monomorphism from A to U. ♢

We have now completed the brief presentation of the notions of AST that will be necessary in Chapter 6. There would be much more to say, as one of the most interesting aspects of AST is the formulation of axioms for small maps related to the choice principles for CST [62]. However, those axioms will not play a role in this thesis, and therefore we prefer to avoid to discuss them.

2.4.2 Categories of classes, concretely

We now give an example of a category $\mathcal{E}$ with a family of small maps that satisfies the properties we just discussed,working informally in **CZF**. The objects of the category will be classes, and hence the whole category will be 'extra-large' in that its objects are neither a set nor a class, but rather a collection of classes. This will be the case also for categories of presheaves that we are to define in Chapter 6. In order to discuss these categories at a completely formal level, one would need to introduce an extension of the language of CST to accommodate variables for classes, and then show that such an extension is conservative. However, we prefer to avoid these tiresome details, and continue the discussion at a rather informal, but hopefully rigorous, level.

Let **Classes** be the category whose objects are classes and maps are function classes, i.e. functional relations between classes. Note that functional relations are generally classes, and not sets. From now on we prefer to write $\mathcal{E}$ instead of **Classes**. The notion of set determines a family of small maps $\mathcal{S}$ in $\mathcal{E}$. For a map F from A to B we say that F is **small** if, for all b in B, the fiber of b along F, i.e. the class

$$\{x \in A \mid Fx = b\}\,,$$

is a set. Observe that a class is a small object in $\mathcal{E}$ if and only if it is a set.

It is straightforward to observe that the category $\mathcal{E}$ is regular and that the other properties required at the start of Subsection 2.4.1 hold. Let us however recall from [81] that $\mathcal{E}$ does not seem to be **exact**, i.e. to have 'well-behaved quotients' of equivalence relations [16]. Given an equivalence relation R on a class A, it seems necessary to assume that R is a set in order to define the quotient of A under R [7, Section 3.4]. This is because quotients are given by the equivalence classes of R, which need not to be sets unless R itself is assumed to be a set.

As for axioms for small maps, the category $\mathcal{E}$ clearly satisfies not only the axioms (**A1**) – (**A6**) in Section C.1, but also the axiom (**P1**) in Section C.2. This is due to the definability of power classes in CST. Remarkably, however, these power classes are not small objects, because of the absence of Power Set. Finally, observe that the class of all sets is clearly a universal object in $\mathcal{E}$.

2.5 Some consequences of the collection axioms

In this section we present some consequences of the collection axioms of **CZF**. The proofs of these results are given in some detail to illustrate the kind of reasoning that is used when working informally in CST. For the moment we work informally in **CZF**$^-$. We will apply these propositions in the development of sheaf interpretations for CST in Part II of the thesis.

Proposition 2.3. *Let a be a set and let ψ be a formula of $\mathcal{L}^{(V)}$. Assume that*

$$(\forall x \in a)(\exists y)\psi\,.$$

Then there exists a function g with domain a such that

$$(\forall x \in a)\big((\exists y)(y \in gx) \wedge (\forall y \in gx)\,\psi\big)\,.$$

Proof. For x, z define

$$\xi =_{\mathrm{def}} (\exists y)(z = (x, y) \wedge \psi)\,.$$

We have $(\forall x \in a)(\exists z)\xi$ by the assumption. By Strong Collection, as given in Section A.4 on page 186 there exists a set u such that

$$\mathrm{coll}(x \in a, z \in u, \xi)\,, \tag{2.3}$$

where we used an abbreviation defined in Section A.4. Define a function g with domain a by letting, for x in a,

$$gx =_{\mathrm{def}} \{y \mid (x, y) \in u\}\,,$$

and observe that g is a set by Replacement. We have the conclusion by (2.3) and the definition of ξ. Discharging the assumption of u, the proof is complete. □

Proposition 2.4. *Let a be a set, let ϕ be a formula of $\mathcal{L}^{(V)}$ and let Q be a class. Assume that*

$$(\forall x \in a)(\exists y)\big(y \subseteq Q_x \wedge \phi\big) \wedge (\forall x \in a)(\forall y)(\forall z)\big((y \subseteq z \subseteq Q_x \wedge \phi) \rightarrow \phi[z/y]\big)$$

where, for x in a, $Q_x =_{\mathrm{def}} \{y \mid (x, y) \in Q\}$. There exists a function f with domain a such that

$$(\forall x \in a)(fx \subseteq Q_x \;\wedge\; \phi[fx/y]).$$

Proof. For x, y define

$$\psi =_{\mathrm{def}} y \subseteq Q_x \wedge \phi$$

We have $(\forall x \in a)(\exists y)\psi$ by the assumption. There exists a function g with domain a such that

$$(\forall x \in a)\big((\exists y)(y \in gx) \wedge (\forall y \in gx)\psi\big) \tag{2.4}$$

by Proposition 2.3. Define a function f with domain a by letting, for x in a,

$$fx =_{\mathrm{def}} \bigcup gx\,,$$

i.e. $(\forall z)(z \in fx \leftrightarrow (\exists y \in gx) z \in y)$, and observe that f is a set by Union and Replacement. For x in a we now show

$$fx \subseteq Q_x \wedge \phi[fx/y]$$

To prove the first conjunct, let z in fx. There exists y in gx such that $z \in y$ by the definition of f. We have $y \subseteq Q_x$ by (2.4) and the definition of ψ. Hence $z \in Q_x$. Discharging the assumption of y we have $fx \subseteq Q_x$, as wanted. To prove the second conjunct, observe that there exists y in gx such that ψ by (2.4). We have

$$y \subseteq fx \subseteq Q_x \wedge \phi$$

by the definitions of f and ψ. Therefore $\phi[fx/y]$ by the assumption in the statement of the proposition. Discharging the assumption of y, we obtain the desired conclusion. We have therefore proved the two conjuncts. Universally quantifying over x and discharging the assumption of g, the proof is complete. □

Proposition 2.5. *Let a be a set, let ϕ be a formula of $\mathcal{L}^{(V)}$ and let P be a class. Assume that*

$$(\forall x \in a)\Big((\exists y)(y \subseteq P \wedge \phi) \wedge (\forall y)(\forall z)\big((y \subseteq z \subseteq P \wedge \phi) \rightarrow \phi[z/y]\big)\Big)\,.$$

Then there exists a set b such that

$$b \subseteq P \wedge (\forall x \in a)\phi[b/y].$$

Proof. Define

$$Q =_{\mathrm{def}} \{(x, y) \mid x \in a \wedge y \in P\}$$

For all x in A we have $Q_x = P$, where Q_x is defined as in Proposition 2.4. We have

$$(\forall x \in a)(\exists y)\big(y \subseteq Q_x \wedge \phi\big) \wedge (\forall x \in a)(\forall y)(\forall z)\big((y \subseteq z \subseteq Q_x \wedge \phi) \rightarrow \phi[z/y]\big)$$

by the assumption. Therefore there exists a function g with domain a such that

$$(\forall x \in a)(fx \subseteq P \wedge \phi[fx/y]) \tag{2.5}$$

by Proposition 2.4 and the definition of Q. Define

$$b =_{\text{def}} \bigcup_{x \in a} fx \,,$$

i.e. $(\forall y)(y \in b \leftrightarrow (\exists x \in a) y \in fx)$ and observe that b is a set by Union and Replacement. We have $b \subseteq P$ by the definition of f and (2.5). Let x in a, and observe that

$$fx \subseteq b \subseteq A \wedge \phi[fx/y]$$

by the definition of b and (2.5). Therefore $\phi[b/y]$ by the assumption in the statement of the proposition. Universally quantifying over x and discharging the assumption of f, the proof is complete. □

The next result was obtained in [1] but we present a proof for completeness. We now work informally in **CZF** and use Subset Collection.

Proposition 2.6. *Let a and b be sets. Let ϕ be a formula. Then there exists a set c such that*

$$(\forall u \in a)\,(\forall z)\big((\forall x \in u)(\exists y \in b)\phi \rightarrow (\exists v \in c)\,\text{coll}\,(x \in u, y \in v, \phi)\big)$$

Proof. For u and w define

$$\psi =_{\text{def}} (\forall z)\big((\forall x \in u)(\exists y \in b)\phi \rightarrow (\exists v \in w)\text{coll}(x \in u, y \in v, \phi)\big)$$

We have $(\forall u \in a)(\exists w)\psi$ by Subset Collection, as given in Section A.4. Therefore there is d such that

$$\text{coll}(u \in a, w \in d, \psi) \tag{2.6}$$

by Strong Collection. Define $c =_{\text{def}} \bigcup d$, i.e.

$$(\forall v)(v \in c \leftrightarrow (\exists w \in d) v \in w)\,.$$

We now show that c satisfies the conclusion. Let u in a. Let z be a set. Assume

$$(\forall x \in u)(\exists y \in b)\phi$$

We have that there is w in d such that ψ by (2.6). Then there is v in w such that

$$\text{coll}(x \in u, y \in v, \phi)$$

by definition of ψ and the assumption. Hence the conclusion. Discharging the assumption of d the proof is complete. □

To simplify some of the applications of Proposition 2.6 that are used in the following, we will depart from the strict application of the laws of logic, and rather adopt an abuse that we now justify. For sets u, b, z and a formula ϕ we will often prove a statement of the form

$$(\forall x \in u)(\exists y \in b)\,.$$

In this circumstances, we will claim that there exists a set c independent of u, b, and z for which there is an element v of c such that

$$\mathrm{coll}\,(x \in u, y \in v, \phi)$$

holds. This is clearly justified by Proposition 2.6, provided that the sets u are elements of a set a, as it will always be the case.

Chapter 3

Pure type theories

3.1 Judgements and raw syntax

In this chapter we present pure type theories and fix some notation and terminology related to them. This will be convenient when, in Chapter 7, we introduce logic-enriched type theories as extensions of pure type theories. Given the considerable introductory literature on DTT, our presentation of pure type theories will be very brief, and we invite the reader who is not familiar with DTT to refer to [59, 65] for more information.

3.1.1 Judgements

A **standard pure type theory** has the forms of judgement $(\Gamma)\ \mathcal{B}$ where Γ is a context consisting of of a list of declarations $x_1 : A_1, \ldots, x_n : A_n$ of distinct variables $x_1, \ldots, x_n$, and $\mathcal{B}$ is a **judgement body** of one of the following forms

$$\begin{aligned} A &: \mathsf{type}\,, \\ A_1 = A_2 &: \mathsf{type}\,, \\ a &: A \\ a_1 = a_2 &: A\,. \end{aligned}$$

All the type theories we will consider are standard, and therefore from now on we prefer to say 'pure type theory' rather than 'standard pure type theory' for brevity.

Let us now introduce some notions related to the syntactic formulation of pure type theories. We say that a context Γ consisting of declarations $x_1 : A_1, \ldots, x_n : A_n$

is **well-formed** if the judgements

$$\begin{array}{rcl} (\ )\ A_1 & : & \mathsf{type}\,, \\ (x_1 : A_1)\ A_2 & : & \mathsf{type}\,, \\ \cdots & & \cdots\,, \\ (x_1 : A_1, \ldots, x_{n-1} : A_{n-1})\ A_n & : & \mathsf{type}\,, \end{array}$$

are derivable. The well-formedness of each of the forms of judgement that are part of a pure type theory has presuppositions other than the well-formedness of its context. In a well-formed context Γ, the judgement $A_1 = A_2 : \mathsf{type}$ presupposes that $A_1 : \mathsf{type}$ and $A_2 : \mathsf{type}$; the judgement $a : A$ presupposes that $A : \mathsf{type}$; finally, the judgement $a_1 = a_2 : A$ presupposes that $a_1 : A$ and $a_2 : A$.

A convention. In the rest of the thesis we prefer to leave out the empty context whenever possible. So $(\ )\ A : \mathsf{type}$ will be written just $A : \mathsf{type}$.

Any standard type theory will have **rules** for deriving well-formed judgements, each instance of a rule having the form

$$\frac{J_1 \quad \cdots \quad J_k}{J}$$

where $J_1, \ldots, J_k, J$ are all judgements.

Notation. In stating a rule of a standard type theory it is convenient to suppress mention of a context that is common to both the premisses and the conclusion of the rule. For example we will write the reflexivity rule for type equality as just

$$\frac{A : \mathsf{type}}{A = A : \mathsf{type}\,,}$$

but in applying this rule we are allowed to infer $(\Gamma)\ A = A : \mathsf{type}$ from $(\Gamma)\ A : \mathsf{type}$ for any well-formed context Γ.

3.1.2 Raw syntax

When it comes to define interpretations of type theories, it is convenient to have a systematic account of their syntax. Here, we prefer to specify a raw syntax rather than introducing a logical framework [65]. This will only be used in Section 8.6, where we discuss types-as-classes interpretations. The raw syntax we adopt categorises expressions into two groups of:

- individual expression (i.e. term),
- type expression.

It is important to observe that raw expressions need not be well-formed expressions of a type theory. We will refer to terms and type expressions as **0-expressions** and **1-expressions** respectively. The raw expressions will be built up from an unlimited supply of individual variables and a signature of constant symbols according to the rules given below. We assume that each constant symbol of the signature has been assigned an arity $(n_1^{\epsilon_1} \cdots n_k^{\epsilon_k})^{\epsilon}$ where $k \geq 0$, $n_1, \ldots, n_k \geq 0$ and each of $\epsilon, \epsilon_1, \ldots, \epsilon_k$ is either 0 or 1. We say that a symbol of such an arity is k-**place**. The rules for forming raw expressions of the two kinds are as follows:

- every variable is a 0-expression,
- if κ is a constant symbol of arity $(n_1^{\epsilon_1} \cdots n_k^{\epsilon_k})^{\epsilon}$ and M_i is an ϵ_i-expression and $\vec{x}_i$ is a list of n_i distinct variables, for $i = 1, \ldots, k$, then

$$\kappa((\vec{x}_1)M_1, \ldots, (\vec{x}_k)M_k) \tag{3.1}$$

is an ϵ-expression.

Some conventions. For a symbol κ as in (3.1) we just write κ rather than $\kappa(\ )$ if $k = 0$. Also, if $n_i = 0$ for some i then we write just M_i rather than $(\ )M_i$.

Free and bound occurrences. These are defined in the standard way when the $(\vec{x}_i)$ are treated as variable binding operations, so that free occurrences in M_i of variables from the list $\vec{x}_i$ become bound in $(\vec{x}_i)M_i$ and so they are bound also in the whole expression $\kappa((\vec{x}_1)M_1, \ldots, (\vec{x}_k)M_k)$.

Substitution. The result $M[M_1, \ldots, M_k/y_1, \ldots, y_k]$ of simultaneously substituting M_i for free occurrences of y_i in M for $i = 1, \ldots, k$, where $y_1, \ldots, y_k$ are distinct variables, is defined in the usual way, relabeling bound variables as usual so as to avoid variable clashes. This is only uniquely specified up to α-convertibility, i.e. up to suitable relabeling of bound variables. In general expressions will be identified up to α-convertibility.

Recall that raw expressions need not be well-formed expressions; it is indeed the two judgement forms

- $(\Gamma)\ a : A$,

- $(\Gamma)\ A : \mathsf{type}$,

that are used to express that

- a is a well-formed term of type A ,

- A is a well-formed type,

in the context Γ.

3.2 Rules for pure type theories

A type theory is specified by indicating a signature for constant symbols and rules to derive judgements. In the following we prefer to leave implicit the signature of a type theory, and just indicate its rules. The rules for all the type theories that we will consider in the following are explicitly presented in Appendix B. The rule for a pure type theory can be divided into two groups: general rules, listed in Section B.1 and type rules, presented in Section B.2.

General rules can in turn be divided into assumption, equality, substitution and congruence rules. Assumption rules specify how to extend a context with additional variable declarations. Equality rules regard the judgement forms

$$\begin{array}{lcll} A_1 & = & A_2 & : \mathsf{type}\,, \\ a_1 & = & a_2 & : A\,. \end{array} \tag{3.2}$$

They serve two purposes: on the one hand, they express that both equality between types and equality between terms are equivalence relations; on the other hand, they specify how the two judgements forms in (3.2) interact with each other. Substitution rules indicate the behaviour of substitution in contexts and in judgement bodies. Finally, congruence rules specify the interplay between the equality on types and the substitution and type formation rules.

Type rules can be divided into groups according to the form of type they regard. For each form of type there will be formation, introduction, elimination and computation rules. The different roles played by these groups of rules is explained, for example, in [59, pages 24 – 25]. We say that a term expression is **canonical** if has the same form of a type expression appearing in the conclusion of an introduction rule.

3.3 A review of pure type theories

A variety of pure type theories can be determined by alternative selections of rules. In this section we review the formulation of the pure type theories that will be considered in the rest of the thesis.

3.3.1 ML

We use **ML** to stand for a variant of Martin-Löf's type theory without universes or W-types. We prefer to avoid having any identity types. Also, rather than have finite types N_i for all $i = 0, 1, \ldots$ we will just have them for $i = 0, 1, 2$ and indicate them with $\mathsf{0}, \mathsf{1}$ and $\mathsf{2}$, respectively. We also have a type of natural numbers, that is indicated by N. These are the non-dependent types of **ML**, whose rules are presented in Subsection B.2.1.

We do not take binary sums as primitive, but rather define them. In order to do so, we allow dependent types to be defined by cases on the type $\mathsf{2}$, i.e. we have the R_2-formation rule

$$\frac{A_1 : \mathsf{type} \qquad A_2 : \mathsf{type} \qquad e : \mathsf{2}}{\mathsf{R}_2(A_1, A_2, e) : \mathsf{type}}$$

and the R_2-equality rules

$$\begin{aligned} \mathsf{R}_2(A_1, A_2, 1_2) &= A_1 \,, \\ \mathsf{R}_2(A_1, A_2, 2_2) &= A_2 \,, \end{aligned}$$

where 1_2 and 2_2 are the canonical elements of the type $\mathsf{2}$. Finally, we have Σ-types and Π-types. Define, for $e : (\Sigma x : A)B$, the first and second projection of e as

$$\begin{aligned} e.1 &=_{\mathrm{def}} \mathsf{split}((x, y)x, e) \,, \\ e.2 &=_{\mathrm{def}} \mathsf{split}((x, y)y, e) \,. \end{aligned}$$

Observe that with these definitions, the rules

$$\frac{e : (\Sigma x : A)B}{e.1 : A} \qquad\qquad \frac{e : (\Sigma x : A)B}{e.2 : B[e.1/x]}$$

are derivable. Therefore the primitive forms of type of **ML** are

$$\mathsf{0}\,,\ \mathsf{1}\,,\ \mathsf{2}\,,\ \mathsf{N}\,,\ \mathsf{R}_2(A_1, A_2, e)\,,\ (\Sigma x : A)B\,,\ (\Pi x : A)B\,.$$

The rules for the dependent types of **ML** are given in Subsection B.2.2. Having these forms of type, we can define others. For $A_1, A_2 : \mathsf{type}$ define the product, function, and binary sum forms of type respectively as

$$\begin{aligned} A_1 \times A_2 &=_{\text{def}} (\Sigma\_ : A_1)A_2\,, \\ A_1 \to A_2 &=_{\text{def}} (\Pi\_ : A_1)A_2\,, \\ A_1 + A_2 &=_{\text{def}} (\Sigma z : 2)\mathsf{R}_2(A_1, A_2, z)\,, \end{aligned}$$

where the symbol _ indicates an anonymous bound variable. Special instances of Σ-rules and Π-rules allow to derive the familiar rules concerning product and function types. Furthermore, the combination of R_2-rules and Σ-rules allows us to derive the rules for binary sum types.

3.3.2 $\mathbf{ML_1}$

The type theory $\mathbf{ML_1}$ is obtained from **ML** by adding the U-rules for a type universe, as presented in Subsection B.2.3. In particular, we have the elimination rule

$$\frac{a : \mathsf{U}}{\mathsf{T}\, a : \mathsf{type}}\,,$$

and hence our formulation follows the so-called 'Tarski style' [59, pages 88 – 89]. For $A : \mathsf{type}$ we say that A is a **small** if, for some $a : \mathsf{U}$, the judgement

$$A = \mathsf{T}\, a : \mathsf{type}$$

is derivable, and in that case we say that a is as a **representative** for A. In $\mathbf{ML_1}$ the U-rules express that the type universe U reflects all the forms of type of **ML**.

3.3.3 MLW

We may consider adding W-rules for types of well-founded trees, as in Subsection B.2.2, to **ML**. We write **MLW** for the resulting type theory. When combining W-rules and U-rules we may choose whether to have rules to reflect W-types in the type universe. If we do, then the type theory we obtain is written as $\mathbf{MLW_1}$. This type theory has two natural subtheories: the first one, called $\mathbf{ML_1} + \mathbf{W}$, has arbitrary W-types, but they are not reflected in the type universe; the second one, denoted as $\mathbf{ML_{1W}}$, has only small W-types and they are reflected in the type universe.

Remark. Observe that the type theory that we write here as $\mathbf{MLW_1}$, has been denoted in [40] with a different notation. The notation used there seems a bit misleading as it suggests that the universe does not reflect the W-types.

3.3.4 $\mathbf{ML_1^-} + \mathbf{W}^-$

We now introduce a pure type theory that will be of particular interest in connection to constructive set theories. This type theory is a subsystem of $\mathbf{ML_1} + \mathbf{W}$ and it is obtained by restricting the rules for Π-types and W-types. To explain the restrictions, let us work informally in $\mathbf{ML_1} + \mathbf{W}$. For $a : \mathsf{U}$ and $(x : \mathsf{T}\, a)\ B : \mathsf{type}$ define

$$(\Pi^- x : a)B =_{\text{def}} (\Pi x : \mathsf{T}\, a)B : \mathsf{type}\,, \tag{3.3}$$

and observe that the rule

$$\frac{a : \mathsf{U} \qquad (x : \mathsf{T}\, a)\ B : \mathsf{type}}{(\Pi^- x : a)B : \mathsf{type}}$$

is derivable. Similarly, for $A : \mathsf{type}$ and $(x : A)\ b : \mathsf{U}$ define

$$(W^- x : A)b =_{\text{def}} (W x : A)\, \mathsf{T}\, b\,, \tag{3.4}$$

and observe that the rule

$$\frac{A : \mathsf{type} \qquad (x : A)\ b : \mathsf{U}}{(W^- x : A)b : \mathsf{type}}$$

is derivable. We may easily formulate introduction, elimination and computation rules associated for the forms of type defined in (3.3) and (3.4). These rules can obviously be derived in $\mathbf{ML_1} + \mathbf{W}$, but we may also consider a type theory that has these restricted rules instead of the more general ones, as we do now. The type theory $\mathbf{ML_1^-} + \mathbf{W}^-$ in obtained from $\mathbf{ML_1} + \mathbf{W}$ by replacing Π-rules and W-rules with Π^--rules and W^--rules, as given in Section B.3.

One of the reasons for the interest in this pure type theory is its close connection to constructive set theories. In Part III we will indeed show that it has a logic-enriched extension $\mathbf{ML(CZF)}$ which is mutually interpretable with the constructive set theory $\mathbf{CZF}$.

Remark. The reader is invited to observe that $\mathbf{ML_1^-} + \mathbf{W}^-$ does not have any U-rule reflecting W^--types because $\mathbf{ML_1} + \mathbf{W}$ does not have any U-rule reflecting W-types. It is also worth pointing out that $(\Pi^- x : a)B$ and $(W^- x : A)b$ are primitive expressions in $\mathbf{ML_1^-} + \mathbf{W}^-$, while they are defined ones in $\mathbf{ML_1} + \mathbf{W}$.

Part II

Sheaf interpretations for CST

Chapter 4

Exercises in formal topology

4.1 Introduction

Even without considering its interaction with intuitionistic mathematics, the field of **pointfree topology** stands out as remarkable. At first, the idea of developing topology without assuming the notion of point as primitive may seem rather eccentric. The vast literature on the subject proves however that such an approach is not only effectively possible, but mathematically efficient [48]. If we then consider the relationship of pointfree topology with intuitionistic mathematics, the subject reveals itself to be even more fruitful since it leads to a deeper insight into the connections between intuitionistic logic and topology [23, 30] and allows a number of applications [31]. The present chapter explores some aspects of this connection.

Pointfree topology has a twofold interaction with intuitionistic mathematics at the fully impredicative level. The first aspect of this interplay stems from the observation that both the internal logic of toposes and intuitionistic set theories provide a suitable setting for the development of much of pointfree topology [47]. The second aspect of the interplay arises since the notion of frame, that is central to pointfree topology, determines two important classes of toposes and of interpretations for intuitionistic set theories: localic toposes [56, Chapter IX] and Heyting-valued interpretations [38]. Both aspects are essential to the relationship between pointfree topology and intuitionistic mathematics: the first one illustrates that toposes and intuitionistic set theories are systems capable of supporting the development of substantial parts of mathematics, while the second one provides a wide range of applications for pointfree topology.

The field of **formal topology** originated by considering the first aspect of the interaction between pointfree topology and intuitionistic mathematics, but at the generalised predicative level [68]. One of the original aims of formal topology is indeed to investigate whether pointfree topology could be developed within Martin-Löf pure type theories. Considerable research indicates that this is possible, at least for parts of pointfree topology, and that this undertaking also leads to the formulation of notions that might be of interest at the fully impredicative level [69]. Yet, the second aspect of interplay between pointfree topology and intuitionistic mathematics does not seem to have been explored at the generalised predicative level. One of the historical reasons for this fact seems to be that formal topology has been traditionally developed in the setting of pure type theories [71], where Heyting-valued interpretations have not been explored.

We set out to explore formal topology in Constructive Set Theory (CST) and investigate whether it is possible that an interplay exists between formal topology and CST of the kind existing between pointfree topology and Intuitionistic Set Theory (IST). The present chapter investigates whether constructive set theories provide an appropriate setting for the development of formal topology, while the next chapter studies the possibility of developing Heyting-valued interpretations for CST.

There are at least two main motivations for focusing on constructive set theories: the first one is that the familiar set theoretic notation allows a rather straightforward development of intuitionistic mathematics, and the second one is that the considerable research on Heyting-valued interpretations for intuitionistic set theories suggests that they may carry over to constructive set theories [31, 38, 72, 73]. A further motivation for the work described in this chapter is that formal topology has traditionally been developed assuming the propositions-as-types treatment of logic [71]. Although this treatment of logic is believed not to be necessary [69], the possibility of avoiding it does not seem to have been explicitly explored yet. It seems therefore natural to test whether the axioms of CST are sufficient to support the development of formal topology.

In spite of the familiar set theoretic notation that is available in CST, this undertaking faces a number of challenges because of the specific axioms that are assumed in CST. For example, the very notion of frame needs to be reconsidered, since the absence of Power Set and Full Separation prevents us from showing that many familiar examples of frames in IST are sets in CST. The paradigmatic example of this situation, i.e. the power class of $\mathbf{1}$, was discussed in Section 2.3.

We will therefore introduce the notions of **set-generated** and **set-presented frame**, that are based on the notions of set-generated and set-presented $\bigvee$-semilattice introduced in [7, Chapter 6]. We hasten to say that these notions are equivalent to notions of formal topology existing in the literature [68, 69] and therefore their introduction does not constitute a novelty. Set-generated and set-presented frames, however, allow us to start developing formal topology in CST in a very direct way, following very closely the development of pointfree topology in intuitionistic set theories, but assuming only axioms of constructive set theories. The peculiar feature of set-generated frames is that, while their carriers are allowed to be proper classes, we require the existence a set of 'generators' for the elements of the frame, in the sense that any element of the frame is the supremum of a suitable set of generators.

We only take a few steps in the development of formal topology in CST, essentially as many as it is necessary to set up and apply Heyting-valued interpretations in Chapter 5. First, we introduce set-generated and set-presented frames and show that well-known representation theorems for frames carry over in our context. Secondly, we introduce a notion of point for a set-generated frame and describe how to associate a space of points to any set-generated frame. Thirdly, we show how inductive definitions allow us to transfer to our context the definition of frames by 'generators and relations'. Finally, we present examples for the notions of set-generated and set-presented frames corresponding not only to familiar spaces, such as the Baire and Cantor space, but also to the space of Dedekind reals as defined in CST [7, Section 3.6]. These few indications highlight however that the axioms of CST are sufficient to develop formal topology in CST. For example, the Regular Extension axiom plays a crucial role when it comes to discussing inductive definitions of frames.

4.2 Set-generated frames

In this section we introduce the notion of a set-generated frame, that will be used throughout this chapter to develop formal topology in CST.

4.2.1 Set-generated $\bigvee$-semilattices

Let us recall from [7, Chapter 6] the notion of set-generated $\bigvee$-semilattice and provide some motivation for it. The first step to introduce this notion is to consider a partially ordered structure whose carrier may be a class, as we do next. Let A be a class and R be a relation on A, i.e. a subclass of $A \times A$. For a, b in A, we write $R(a, b)$ instead of $(a, b) \in R$. We say that R is a **partial order** on A if it holds that

- R is reflexive, i.e. $(\forall x \in A)R(x, x)$,
- R is transitive, i.e. $(\forall x, y, z \in A)R(x, y) \wedge R(y, z) \rightarrow R(x, y)$,
- R is antisymmetric, i.e. $(\forall x, y \in A)R(x, y) \wedge R(y, x) \rightarrow x = y$.

If R is a partial order on A we write $a \leq b$ instead of $R(a, b)$ for a, b in A .

Definition 4.1. We say that (A, R) is a partially ordered class, or **poclass** for short, if A is a class and R is a partial order on it. We say that a poclass (A, R) is a partially ordered set, or **poset** for short, if both A and R are sets. $\Diamond$

Remark. In the following, we sometimes prefer to avoid reference to all the structure that is formally part of a poclass and simply say 'A is a poclass', leaving implicit the partial order that is part of the poclass structure. We will use an analogous convention for the other structures that we introduce in the remainder of this chapter.

Using this convention, we now introduce the notion of morphism of poclasses. Let A and B be poclasses, we say that a function ϕ from A to B is a **morphism** of poclasses if it holds that

- ϕ is monotone, i.e. $(\forall x, y \in A) x \leq y \rightarrow \phi(x) \leq \phi(y)$.

We say two poclasses are **isomorphic** if there are two mutually inverse morphisms between them. Poclasses will be the structure on which all the other notions introduce in this section are based. When considering additional structure on a poclass, the distinction between subsets and subclasses that is peculiar to CST plays an important role, that we will often point out to the reader. The notion of supremum of a subclass, that we introduce next, is a typical example of this phenomenon. Let A be a poclass and let P be a subclass of A. We say that an element a of A is the **supremum** of P if it holds that

$$(\forall x \in A) \big(a \leq x \leftrightarrow (\forall y \in P) y \leq x\big) .$$

Our use of the definite article 'the' when introducing the notion of supremum is justified because if a supremum exists then it is unique. Similar considerations apply when we will introduce the notions of bottom, infimum, top, meet and Heyting implication. In the following we will write $\bigvee P$ for the supremum of a subclass P, if it exists.

Definition 4.2. We say that $(A, \leq, \bigvee)$ is a **$\bigvee$-semilattice** if $(A, \leq)$ is a poclass and $\bigvee$ is a supremum operation for it, i.e. an operation assigning to each subset of A its supremum. ◊

Remark. Let us stress that in a $\bigvee$-semilattice the supremum operation is assumed to act only on subsets and not on subclasses. However, the language of set theory allows us to say what it means for an element to be the supremum of a subclass.

Let A be a poclass. We say that a in A is the **bottom element** in A if it holds that $(\forall x \in A) a \leq x$. We will write $\perp$ for the bottom element of a poclass, when it exists. It is a simple observation to note that every $\bigvee$-semilattice has a bottom element, that can be defined as the supremum of the empty set.

The notion of infimum, that we introduce next, will help us to motivate the notion of generating set. We say that an element a of A is the **infimum** of P if it holds that

$$(\forall x \in A)\big(x \leq a \leftrightarrow (\forall y \in P)x \leq y\big)\,.$$

A **complete lattice** is a $\bigvee$-semilattice with an infimum operation, i.e. an operation assigning to each subset of the $\bigvee$-semilattice its infimum. A remarkable fact distinguishes the development of lattice theory based on posets, as done in IST, from the one based on poclasses, as done here. It is well-known that if A is a $\bigvee$-semilattice that is a set, then it is also a complete lattice [46, Section 4.3]. It is worth, however, highlighting that the assumption that A is a set plays a role to prove this simple fact. Let p is a subset of A and define

$$q =_{\mathrm{def}} \{x \in A \mid (\forall y \in p)x \leq y\}\,.$$

Since A and p are sets, so is q. We can then observe that the infimum is given by the supremum of q, which exists since q is a set and A is a $\bigvee$-semilattice. This reasoning indicates that it is not possible to define an infimum operation on an arbitrary $\bigvee$-semilattice without further assumptions. This problem can be overcome introducing the following notion. Let A be a $\bigvee$-semilattice and g be a subset of A; for a in A define

$$g_a =_{\mathrm{def}} \{x \in g \mid x \leq a\}\,.$$

We say that g is a **generating set** for A if the following properties

- for all a in A the class g_a is a set,
- for all a in A, $a = \bigvee g_a$,

hold. As we will see, the assumption of the existence of a generating set for A is more general than the assumption that A is a set, and it is sufficient to define an infimum operation.

Definition 4.3. We say that $(A, \leq, \bigvee, g)$ is a **set-generated $\bigvee$-semilattice** if $(A, \leq, \bigvee)$ is a $\bigvee$-semilattice and g is a generating set for it. ♢

Remark. We invite the reader to note that the generating set is formally part of the structure of a set-generated $\bigvee$-semilattice. Although the convention we fixed in the remark after Definition 4.1 allows us to say 'A is a set-generated $\bigvee$-semilattice', it is sometimes necessary to specify the generating set for A. We will do so by saying 'A is a set-generated $\bigvee$-semilattice with generating set g'.

Let A and B be set-generated $\bigvee$-semilattices. We say that a function ϕ from A to B is a **morphism** of set-generated of $\bigvee$-semilattices if it holds that

- ϕ is monotone,

- ϕ preserves suprema, i.e. for all subsets p of A, $\phi(\bigvee p) = \bigvee\{\phi(x) \mid x \in p\}$.

Note that the generating sets do not play a role in the properties of morphisms.

Proposition 4.4. *Every set-generated $\bigvee$-semilattice is a complete lattice.*

Proof. Let A be a set-generated $\bigvee$-semilattice with generating set g. For p in $\mathsf{Pow}\,A$, define $q =_{\text{def}} \{x \in g \mid (\forall y \in p) x \leq y\}$ and observe that q is a set by the assumption that g is a generating set and Replacement. Define $\bigwedge p =_{\text{def}} \bigvee q$ and observe that $\bigwedge p$ is an infimum for p. □

Let A be a poclass. We say that a in A is the **top element** of A if it holds that $(\forall x \in A) x \leq a$. From now on we write $\top$ for the top element of a poclass. For a and b in A, we say that an element c in A is the **meet** of a and b if it holds that

$$(\forall x)\big(x \leq c \leftrightarrow (x \leq a \wedge x \leq b)\big)\,.$$

We write $a \wedge b$ for the meet of a and b if it exists. It is immediate show that any complete lattice, and hence any set-generated $\bigvee$-semilattice, has a top element and a meet operation: the top element is defined as the infimum of the empty set, and the meet of two elements is the infimum of the set consisting only of them. We now give some examples of set-generated $\bigvee$-semilattices.

Lower sets. Let s be a poset. For a subclass P of s define

$$\delta P =_{\text{def}} \{x \in s \mid (\exists y \in P) x \leq y\}$$

and observe that $P \subseteq \delta P$. We say that P is a **lower class** if it holds that $\delta P = P$, and say that it is a **lower set** if it is also a set. Define $\mathsf{Low}(s)$ to be the poclass of lower sets of s, with partial order given by inclusion. The supremum operation is union, since the union of a set of lower sets is a lower set. In order to show that $\mathsf{Low}(s)$ has a generating set it is convenient to define, for a in s, $\gamma a =_{\text{def}} \delta\{a\}$. Using this abbreviation, define

$$g =_{\text{def}} \{\gamma x \mid x \in s\}$$

and observe that g is a generating set for $\mathsf{Low}(s)$. Hence $\mathsf{Low}(s)$ is a set-generated $\bigvee$-semilattice. The bottom element is the empty set. Direct calculations lead to show that if u is a subset of $\mathsf{Low}(s)$ then

$$\bigwedge u = \bigcap u\,.$$

The top element of the set-generated $\bigvee$-semilattice is s, which is a lower set.

Subsets. Any set s can be seen as a poset by considering the partial order given by equality. In this case lower sets are just subsets, and therefore we obtain that $\mathsf{Pow}(s)$ is a set-generated $\bigvee$-semilattice. Its generating set has a particularly simple form: for a in s define $\gamma a =_{\text{def}} \{a\}$ and define

$$g =_{\text{def}} \{\gamma x \mid x \in s\}\,.$$

Now observe that g is a generating set for $\mathsf{Pow}(s)$.

When considering poclasses with structure it is convenient to have some general methods to define them. The notion of closure operator, that we introduce next, provides a way to define set-generated $\bigvee$-semilattices.

Definition 4.5. Let A be a set-generated $\bigvee$-semilattice and let c be a function from A to A. We say that c is a **closure operator** if the following properties

- c is inflationary, i.e. $(\forall x \in A) x \leq c(x)$,
- c is monotone,
- c is idempotent, i.e. $(\forall x \in A) c(cx) \leq c(x)$,

hold. $\Diamond$

The next definition will be convenient to state Proposition 4.6. For a set-generated $\bigvee$-semilattice A and a closure operator c on it, define

$$A_c =_{\text{def}} \{x \in A \mid x = cx\}\,.$$

The next result is the reformulation in our context of well-known facts in formal and pointfree topology [9, 46], and it is a slight variation over Theorem 6.3 of [7]. Our formulation makes use of lower sets rather than arbitrary subsets, and will therefore link up in a natural way with the treatment of set-generated frames using nuclei in Subsection 4.2.2. In turn, lower sets and nuclei will provide an intuitive connection with the development of presheaf and sheaf interpretations, that are to be discussed in Chapter 6, where we consider Lawvere-Tierney operators.

Proposition 4.6. *Let A be a set-generated $\bigvee$-semilattice with generating set g.*

(i) *If c is a closure operator on A, then A_c determines a set-generated $\bigvee$-semilattice.*

(ii) *There exists a poset s and a closure operator c on $\mathsf{Low}(s)$ such that A and $(\mathsf{Low}\, s)_c$ are isomorphic.*

Proof. For part (i) let us define a supremum operation $\bigvee_c$ on A_c by letting, for a subset p of A_c,

$$\bigvee_c p =_{\text{def}} c\big(\bigvee p\big)\,,$$

where the supremum operation on the right-hand side of the definition is the one of A. For a in g define $\gamma_c(a) =_{\text{def}} c(a)$ and observe that

$$\{\gamma_c(x) \mid x \in g\}$$

is a generating set for A_c. Simple calculations lead to show that we have just defined the structure of a set-generated $\bigvee$-semilattice on A_c, as required. Let us now indicate the proof of part (ii). Firstly, define s to be the generating set of A. Secondly, for p in $\mathsf{Low}(s)$, define

$$cp =_{\text{def}} \{x \in s \mid x \leq \bigvee p\} \tag{4.1}$$

and observe that c is a closure operator on $\mathsf{Low}(s)$, and hence $(\mathsf{Low}\, s)_c$ is a set-generated frame by part (i). The desired conclusion follows by the assumption that A is set-generated. □

Observe that the definition in (4.1) makes sense for arbitrary subsets of s and not just for the lower ones. Hence c can be extended to a closure operator on $\mathsf{Pow}(s)$. We therefore have two set-generated frames: $(\mathsf{Pow}\, s)_c$ and $(\mathsf{Low}\, s)_c$. Thankfully, they turn out to be identical: the key to prove this is to observe that for all p in $\mathsf{Pow}(s)$ we have

$$cp = c(\delta p)\,.$$

In view of this fact, when we use Proposition 4.6, we will be justified in assuming that the closure operator c extends to a closure operator on $\mathsf{Pow}(s)$, and in avoiding the distinction between $(\mathsf{Pow}\, s)_c$ and $(\mathsf{Low}\, s)_c$ if this is convenient.

4.2.2 Frames

We now move on to introduce the notion of set-generated frame and show how it corresponds to complete Heyting algebras. In order to do so, we need to present the property that characterises frames. Since this law makes sense for complete lattices, we can state it without reference to generating sets. Let A be complete lattice, we say that A satisfies the **frame distributivity law** if, for all elements a of A and all subsets p

of A, the property

$$a \wedge \bigvee p = \bigvee \{a \wedge x \mid x \in p\}$$

holds.

Definition 4.7. We say that a set-generated $\bigvee$-semilattice is a **set-generated frame** if it satisfies the frame distributivity law. $\Diamond$

To define frame morphisms we indicate with a subscript the frame to which a certain operation or distinguished element belongs. Let A and B be set-generated frames. We say that a function ϕ from A to B is a **frame morphism** if it holds that

- ϕ is monotone,
- ϕ preserves top element, i.e. $\top_B = \phi(\top_A)$,
- ϕ preserves meets, i.e. $(\forall x, y \in A)\phi(x) \wedge_B \phi(y) = \phi(x \wedge_A y)$,
- ϕ preserves suprema.

We say that two set-generated frames are **isomorphic** if there are two mutually inverse frame morphisms between them. Observe that two set-generated frames are isomorphic if they are isomorphic as poclasses.

We now come to another result that, however simple, highlights the role of generating sets. We have already seen in Proposition 4.4 that the assumption of a generating set allows us to prove that a $\bigvee$-semilattice is also a complete semilattice. Here we show that generating sets allow us to relate frames to complete Heyting algebras. First of all let us introduce the notion of complete Heyting algebra in our context. Let A be a complete lattice. For a and b in A, an element c of A is the **Heyting implication** of a and b if it holds that

$$(\forall x \in A)\big(x \leq c \leftrightarrow (x \wedge a \leq b)\big) .$$

We will write $a \to b$ for the Heyting implication of a and b, if it exists. A **complete Heyting algebra** is a complete lattice with an Heyting implication operation, i.e. an operation assigning to each pair of elements their Heyting implication. The next result makes use of generating sets to define an Heyting implication in any set-generated frame.

Proposition 4.8. *Every set-generated frame is a complete Heyting algebra.*

Proof. Let A be a set-generated frame. Let g be a generating set for it. For a, b in A define $p =_{\text{def}} \{x \in g \mid x \wedge a \leq b\}$ and observe that p is a set by the assumption that g

is a set. Define $a \to b =_{\text{def}} \bigvee p$. Using the frame distributivity law, we can prove that $a \to b$ is the Heyting implication of a and b. $\square$

An example. The examples of set-generated $\bigvee$-semilattices given in Subsection 4.2.1 are actually examples of set-generated frames, since the frame distributivity law holds in them. This can be easily seen by recalling that unions distribute over intersections. Remembering that we defined $\mathsf{P} =_{\text{def}} \mathsf{Pow}\,1$ in Section 2.3, note that P is a set-generated frame.

The next notion will help us define set-generated frames throughout this chapter and plays the same role for set-generated frames that the notion of closure operator does for set-generated $\bigvee$-semilattices [80].

Definition 4.9. Let A be a set-generated frame. We say that a closure operator j on A is a **nucleus** if the property

- j respects meets, i.e. $(\forall x, y \in A) jx \wedge jy \leq j(x \wedge y)$,

holds. $\Diamond$

The next proposition extends Proposition 4.6 to set-generated frames, and it is completely analogous to well-known results in the literature [9, 46]. Let us use a similar notation to the one used in connection to Proposition 4.6 and, for a set-generated frame A and a nucleus j on it, define

$$A_j =_{\text{def}} \{x \in A \mid x = jx\} .$$

Proposition 4.10. *Let A be a set-generated frame.*

(i) *If j is a nucleus on A, then A_j determines a set-generated frame.*

(ii) *There exists a poset s and a nucleus j on $\mathsf{Low}(s)$ such that A and $(\mathsf{Low}\,s)_j$ are isomorphic.*

Proof. The proof of part (i) consists simply in the observation that the assumption that j preserves meets implies that A_j satisfies the frame distributivity law. For part (ii) observe that the closure operator defined in (4.1) is a nucleus if A satisfies the frame distributivity law. $\square$

We can still assume that a nucleus j as in part (ii) of Proposition 4.10 extends to a closure operator on $\mathsf{Pow}(s)$, but it is important to note that its extension need not be, in general, a nucleus on $\mathsf{Pow}(s)$. This is because j does not preserve meets of arbitrary subsets of s in general, but only of lower ones. The author should confess to have overlooked this issue when stating Theorem 1.7 in [37].

An example. Consider the set-generated frame P and define, for p in P,

$$jp =_{\text{def}} \{x \in \mathsf{1} \mid \neg\neg x \in p\}$$

Easy calculations lead a proof that j is a nucleus. This nucleus, which we refer to as the **double negation nucleus**, will play an important role in the applications of Heyting-valued interpretations in Chapter 5. More examples of nuclei will be presented in Section 4.5.

We end this section by describing explicitly the set-generated frame $(\mathsf{Low}\, s)_j$, as in part (ii) of Proposition 4.10. As already discussed, we can assume that j extends to an operator on $\mathsf{Pow}(s)$ and that for all p in $\mathsf{Pow}(s)$

$$jp = j(\delta p)\,.$$

We begin by recalling the supremum operation of $(\mathsf{Low}\, s)_j$. For a subset u of $(\mathsf{Low}\, s)_j$ observe that

$$\bigvee u = j(\bigcup u)$$

Recall that $g =_{\text{def}} \{\gamma x \mid x \in s\}$ is a generating set for $\mathsf{Low}(s)$, where, for x in s, $\gamma x = \delta\{x\}$. Now observe that the set

$$g_j =_{\text{def}} \{\gamma_j(x) \mid x \in s\}$$

is a generating set for $(\mathsf{Low}\, s)_j$, where, for x in s, $\gamma_j(x) =_{\text{def}} j(\gamma x)$. The other operations, defined in terms of the supremum and the generating set, can be described easily using the fact that, for all p in $(\mathsf{Low}\, s)_j$ we have

$$\bigvee\{\gamma_j(x) \mid x \in p\} = jp\,.$$

We leave the details to the reader and simply recall that, for p and q in $\mathsf{Low}(s)$

$$p \wedge q = p \cap q$$

by an application of the fact that j respects meets. The infimum of a subset u of $\mathsf{Low}(s)$ is given by

$$\bigwedge u = j(\bigcap u)\,.$$

The presence of j is necessary since a nucleus does not necessarily preserve infima.

4.3 Set-presented frames

As we will see in Section 4.5, inductive definitions provide us with a very general method to define set-generated $\bigvee$-semilattices. The question of which set-generated $\bigvee$-semilattices arise in this way was solved by Peter Aczel by introducing the notion of set-presented $\bigvee$-semilattice [7, Section 6.3]. Let us review this notion and its connection to inductive definition of $\bigvee$-semilattices, since this will be relevant in Section 4.5.

4.3.1 Set-presented $\bigvee$-semilattices

Let A be a set-generated $\bigvee$-semilattice with generating set g. Let r be a set relation between elements and subsets of g, i.e. a subset of $g \times \mathsf{Pow}(g)$, and for a in g define $ra =_{\text{def}} \{x \in \mathsf{Pow}\, g \mid (a, x) \in r\}$. We say that r is a **relation set** for A if for all elements a of g and subsets p of g the property

$$a \leq \bigvee p \leftrightarrow (\exists u \in ra) u \subseteq p$$

holds.

Some intuition. Let A be a set-generated $\bigvee$-semilattice with generating set g. For a in g consider the class

$$\{p \in \mathsf{Pow}(g) \mid a \leq \bigvee p\}\,.$$

Elements of this class may be called the 'covers' of a. Now observe that in general this class is not a set. If there exists a relation set for A, however, for many purposes we can replace the class of covers of a with a set of 'generating covers', i.e. ra. Hence one may think of set-generated frames with a relation set as frames for which not only there exists a set of 'generators', but also each generator has a set of 'generating covers'.

Definition 4.11. We say that $\langle A, \leq, \bigvee, g, r\rangle$ is a **set-presented $\bigvee$-semilattice** if $\langle A, \leq, \bigvee\rangle$ is a $\bigvee$-semilattice and (g, r) is a presentation for it, i.e. g is a generating set and r is a relation set for A. $\Diamond$

Example. Let s be a set and recall that $\mathsf{Pow}\, s$ is a set-generated $\bigvee$-semilattice, with generating set

$$g = \{\gamma x \mid x \in s\}\,,$$

where for x in s we defined $\gamma x =_{\text{def}} \{x\}$. Now define

$$r =_{\text{def}} \{(\gamma x, \{\gamma x\}) \mid x \in s\}$$

and observe that r is a relation set for $\mathsf{Pow}(s)$. In a similar way it is possible to show that also set-generated $\bigvee$-semilattices of the form $\mathsf{Low}(s)$ for some poset s are set-presented. We leave the details to the reader.

The next notion will allow us to extend Proposition 4.6 to set-presented frames and to review their relationship with inductive definitions.

Definition 4.12. Let s be a set and c a closure operator on $\mathsf{Pow}(s)$. We say that c is a **set-presented closure operator** if there exists a set r of subsets of s or such that for all a in s and p in $\mathsf{Pow}(s)$ the property

$$a \in cp \leftrightarrow (\exists u \in r)\; a \in cu \wedge u \subseteq p$$

holds. ◊

It will be convenient to have a notion of set-presented closure operator also for $\bigvee$-semilattices of the form $\mathsf{Low}(s)$, for a poset s. The remarks following Proposition 4.6 motivate us to say that such a closure operator is set-presented if it can be extended to a closure operator on $\mathsf{Pow}(s)$ that is set-presented in the sense of Definition 4.12. We can now show how Proposition 4.6 transfers to set-presented $\bigvee$-semilattices.

Proposition 4.13.

(i) *Let s be a poset and let c be a set-presented closure operator on* $\mathsf{Low}(s)$*. Then* $(\mathsf{Low}\, s)_c$ *determines a set-presented $\bigvee$-semilattice.*

(ii) *Let A be a set-presented $\bigvee$-semilattice. There exists a poset s and a set-presented closure operator c on* $\mathsf{Low}(s)$ *such that A is isomorphic to* $(\mathsf{Low}\, s)_c$.

Proof. This is an immediate consequence of Theorem 6.3 of [7] and Proposition 4.6. □

The reader may wonder why in Definition 4.12 the notion of set-presented closure operator has been introduced for closure operators on set-generated $\bigvee$-semilattices of the form $\mathsf{Pow}(s)$, for some set s. The main reason for this is to make the connection with inductive definitions as simple as possible. This relationship will be now explained assuming that the reader has some familiarity with inductive definitions in CST [7, Chapter 5].

Let s be a set and let Φ be an inductive definition on s. Define a function with domain $\mathsf{Pow}(s)$ by defining, for p in $\mathsf{Pow}(s)$,

$$c_\Phi(p) =_{\mathrm{def}} I(\Phi, p)\,, \tag{4.2}$$

where $I(\Phi, p)$ is the smallest Φ-closed subset of s that contains p. Since Φ is an inductive definition, c_Φ is a closure operator [7, Section 6.1]. The key result relating set-presented $\bigvee$-semilattices and inductive definitions consists in the observation that all set-presented closure operators on $\mathsf{Pow}(s)$ are of the form defined in (4.2). The proof of the next result is based on the Set Compactness Theorem for CST, that is proved assuming REA [7, Section 5.5].

Theorem 4.14 (Aczel). *Let s be a set. A function c from $\mathsf{Pow}(s)$ to $\mathsf{Pow}(s)$ is a set-presented closure operator on $\mathsf{Pow}(s)$ if and only if there exists an inductive definition Φ on s such that $c = c_\Phi$.*

Proof. See [7, Section 6.1]. □

4.3.2 Frames

It is straightforward to extend the notions and the results of the previous subsection to set-generated $\bigvee$-semilattices that satisfy the frame distributivity law, i.e. set-generated frames.

Definition 4.15. We say that a set-presented $\bigvee$-semilattice is a **set-presented frame** if it satisfies the frame distributive law. ◊

Similarly to what we did in Subsection 4.2.2, we say that a closure operator on $\mathsf{Low}(s)$, for a poset s, is a **set-presented nucleus** if it is a set-presented closure operator on $\mathsf{Low}(s)$ and it is a nucleus.

Proposition 4.16.

(i) *Let s be a poset and let j be a set-presented nucleus on $\mathsf{Low}(s)$. Then $(\mathsf{Low}\, s)_j$ determines a set-presented frame.*

(ii) *Let A be a set-presented frame. There exists a poset s and a set-presented nucleus j on $\mathsf{Low}(s)$ such that A is isomorphic to $(\mathsf{Low}\, s)_j$.*

Proof. This is a direct consequence of Proposition 4.10 and Proposition 4.13. □

We will present examples of set-presented nuclei in Section 4.6.

4.4 Spaces and points

The notion of point seems to transfer quite smoothly from the fully impredicative to the generalised predicative level. At first, we may indeed consider points of a set-generated frame as frame morphisms into the set-generated frame P defined in Subsection 4.2.2. A problem arises, however, in that frame morphisms are generally classes themselves, and hence cannot be considered as elements of a class. This can be solved by considering as points of a set-generated frame to be the elements of a class that is in bijective correspondence with frame morphisms into P. The elements of this class are essentially the formal points of a formal topology [68].

The problem of accommodating these points as an instance of a suitable notion of space arises next. It should not come as a surprise that the notion of space needs to be carefully considered when working in a generalised predicative setting. Here we will follow the approach of basic pairs [70] with only few modifications, some of which have been suggested by Peter Aczel. These modifications are essential to capture the examples given by considering the points of a frame. Although in many interesting cases it is possible to show that the formal points of a formal topology are a set [22], this is not the case in general. Hence we are led to consider basic pairs in which the concrete points form a class. We first introduce these spaces, and then discuss the points of a set-generated frame.

4.4.1 Concrete pospaces

As already pointed out, we follow [70] with few modifications. The main components of a concrete partially ordered space, or concrete pospace for short, will be a class of points, a class of neighbourhood indices, and a relation between them. The relation between points and indices expresses when a point lies in the neighbourhood associated to an index. It will also be convenient to assume that the class of indices is partially ordered, so as to express inclusion between neighbourhoods. Following the notation of basic pairs, we will write X for the class of points and S for the poclass of indices. For x in X and a in S we will write

$$x \Vdash a$$

if (x, a) is in the subclass of $X \times S$ that is the relation between points and neighbourhood indices. Let us introduce some abbreviations that will be useful when defining the notion of concrete pospace: for x in X define

$$\alpha_x =_{\mathrm{def}} \{a \in S \mid x \Vdash a\} ,$$

and for a in S define

$$B_a =_{\text{def}} \{x \in X \mid x \Vdash a\}\,.$$

We are now ready to introduce concrete pospaces.

Definition 4.17. We say that $(X, \Vdash, S)$ is a **concrete pospace** if the following properties

- S is a poset and for all x in X the class α_x is a set.
- $X = \bigcup_{a \in S} B_a$,
- for all a, b in S, if $a \leq b$ then $B_a \subseteq B_b$,
- for all a, b in S, $B_a \cap B_b \subseteq \bigcup\{B_c \mid c \in S\,,\ c \leq a\,,\ c \leq b\}$,

hold. ◊

We say that a subclass P of X is **open** if it holds that

$$P = \bigcup\{B_a \mid a \in S\,,\ B_a \subseteq P\}\,.$$

For two concrete pospaces, $(X, \Vdash, S)$ and $(X', \Vdash, S')$, we say that a function g from X to X' is **continuous** if for all a' in S' the class

$$g^{-1}(B'_{a'}) =_{\text{def}} \{x \in X \mid g(x) \in B'_{a'}\}$$

is open. We say that two concrete pospaces are **homeomorphic** if there are two continuous functions between them that are mutually inverse. We will present examples of concrete pospaces in Section 4.5. Before this, we show how a set-generated frame has a concrete pospace associated to it.

4.4.2 Points

Recall that in Section 2.3 we defined $\mathsf{P} =_{\text{def}} \mathsf{Pow}(1)$ and illustrated how elements of the set-generated P play the role of truth values in CST. From Subsection 4.2.2 we know also that P is a set-generated frame. It seems natural to consider as points of a set-generated frame the frame morphisms from it to P, just as they are defined as frame morphisms into 2 in pointfree topology. This choice is rather appealing, but not quite the most appropriate: for a set-generated frame A with generating set g, functions from A to P are generally classes and therefore we cannot consider them as elements of a class, which can only be sets. However, a simple observation puts us on the right track.

Recall that for p in P we defined

$$!\,p =_{\text{def}} (\exists\_ \in p)\top\,,$$

and observe that, for a frame morphism ϕ from A to P, if we define

$$F_\phi =_{\text{def}} \{x \in A \mid \,!\,\phi(x)\}\,, \tag{4.3}$$

then the intersection of F_ϕ with the generating set g is a set by Restricted Separation, since for all x in A the formula $!\,\phi(x)$ is restricted. Furthermore, the class F_ϕ and its intersection with g reflect the properties of the frame morphism F. We isolate these properties in the next two definitions, in which we use the symbol & to stand for logical conjunction, in order to avoid confusion with the symbol for a meet operation. We will adopt this symbol, when necessary, also in the remainder of this chapter.

Definition 4.18. Let A be a set-generated frame with generating set g. We say that a subclass F of A is a **set-generated completely prime filter** if

- $F \cap g$ is a set,
- F is inhabited, i.e. $(\exists\_ \in F)\top$,
- F is an upper subclass of A, i.e. $(\forall x, y \in A) x \in F \,\&\, x \leq y \to y \in F$
- F is meet-closed, i.e. $(\forall x, y \in A) x \in F \,\&\, y \in F \to x \wedge y \in F$,
- F is completely prime, i.e. $(\forall u \in \mathsf{Pow}\, A) \bigvee u \in F \to (\exists x \in u) x \in F$,

hold. $\Diamond$

The properties of the intersection of a set-generated completely prime filter with g are characterised in the following notion, that is essentially the notion of formal point of formal topology [68, 69].

Definition 4.19. Let A be a set-generated frame with generating set g. We say that a subset α of g is a **generating filter** if

- α is inhabited,
- α is stable, i.e. $(\forall x, y \in g) x \in \alpha \,\&\, y \in \alpha \to (\exists z \in \alpha) z \leq x \,\&\, z \leq y$,
- α is prime, i.e. $(\forall x \in g)(\forall u \in \mathsf{Pow}\, g) x \in \alpha \,\&\, x \leq \bigvee u \to (\exists y \in u) y \in \alpha$

hold. $\Diamond$

Remark. Observe that if α is a generating filter of a set-generated frame A with generating set g, then α is an upper subset of g, because α is prime.

The next proposition confirms that we formulated the notions appropriately. In the remarks preceding Definition 4.3 we introduced, for a set-generated frame A with generating set g and a in A, the following notation:

$$g_a =_{\text{def}} \{x \in g \mid x \leq a\} \, .$$

This abbreviation will be used frequently in the proof of the next proposition.

Proposition 4.20. *Let A be a set-generated frame with generating set g. There is a bijective correspondence between set-generated completely prime filters of A, generating filters of A, and frame morphisms from A to P.*

Proof. We first define a bijection between set-generated completely prime filters and generating filters. Given a set-generated prime filter F, define

$$\alpha_F =_{\text{def}} F \cap g \, .$$

We first show that α_F is inhabited. Since F is inhabited, let us assume a in F, and observe that we have

$$\bigvee g_a \in F \, ,$$

because A is set-generated, and therefore there exists x in g such that $x \leq a$ and $x \in F$, because F is completely prime. This implies that x is in α_F, since it is both in g and in F. We now prove that α_F is stable, so let x, y in α_F. We have $x \wedge y \in F$, because F is meet-closed and hence

$$\bigvee g_{x \wedge y} \in F$$

because A is set-generated. We derive that there is z in P such that z in g and $z \leq x \wedge y$, since F is completely prime, and therefore we have $z \in \alpha_F$ such that $z \leq x$ and $z \leq y$, as required. To show that α_F is prime it is sufficient to use that F is an upper subclass of A and that it is completely prime.

Let α be a generating filter and define

$$F_\alpha =_{\text{def}} \{a \in A \mid (\exists x \in g_a) x \in \alpha\} \, .$$

It is immediate to observe that $F_\alpha \cap g$ is a set, by Restricted Separation. To show that

F_α is inhabited we can assume that there is a in α, since α is inhabited. We have

$$a \leq \bigvee g_a \,,$$

because A is set-generated, and hence there is x in g_a such that $x \in \alpha$, because α is prime. We thus obtained an element of F_α, which is therefore inhabited.

To show that F_α is an upper subclass of A assume that $a \leq b$ and $a \in F_\alpha$. By definition of F_α there is x in g_a such that $x \in \alpha$. From $x \leq a$ and $a \leq b$ we obtain

$$x \leq \bigvee g_b \,.$$

This fact and the assumption that $x \in \alpha$ imply that there exists $y \in g_b$ such that $y \in \alpha$, because α is prime. But this is exactly showing that $b \in F_\alpha$, as desired. To show that F_α is meet-closed, use that A is set-generated and that α is stable. Finally, to show that F_α is completely prime, let p be a subset of A and assume that

$$\bigvee p \in F_\alpha \,.$$

By definition of F_α there is x in $g_{\bigvee p}$ such that $x \in \alpha$. Since α is prime, we get that there is y in p such that $y \in \alpha$. We now use the fact that α is stable and obtain z in α such that $z \leq x$ and $z \leq y$. Since $z \in g_y$ and $z \in \alpha$, we have found y in p such that $y \in F_\alpha$, as required. A series of routine calculations shows that these two definable operations are mutually inverse.

We now indicate how to obtain a bijection between frame morphisms from A to P and set-generated completely prime filters of A. Given a frame morphism ϕ from A to P, define F_ϕ as in (4.3) on page 62. The assumption that ϕ preserves top, is monotone, preserves meets and preserves suprema leads to straightforward proofs that F_ϕ is inhabited, is an upper subclass, is meet-closed and completely prime, respectively. Vice versa, given F that is a set-generated completely prime filter of A, we can define a function ϕ_F from A to P by letting, for a in A,

$$\phi_F(a) =_{\mathrm{def}} \{\_ \in \mathsf{1} \mid (\exists x \in g_a) x \in F \cap g\} \,.$$

The verification that ϕ is monotone and that it preserves top element and meets is immediate, using the fact that A is set-generated. The assumption that A is set-generated plays a role also in showing that ϕ_F preserves suprema, but this requires some more argument: let p a subset of A, we need to show that

$$\phi_F(\bigvee p) \leq \bigvee \{\phi_F(y) \mid y \in p\} \,.$$

First of all, define

$$g_p =_{\text{def}} \{x \in g \mid (\exists y \in p) x \leq y\} ,$$

and observe that $\bigvee p = \bigvee g_p$ because A is set-generated. Therefore it suffices to show

$$\phi_F(\bigvee g_p) \leq \bigvee \{\phi_F a \mid a \in p\} .$$

Once we assume $!\,\phi_F(\bigvee g_p)$, we have that there is $x \in F$ such that $x \leq \bigvee g_p$ by the definition of ϕ_F. Since $x \in F$, we have that there is y in g_p such that $y \in F$, because F is an upper subclass and it is completely prime. Simple calculations now lead us to show that there is z in p such that $!\,\phi_F(z)$ and therefore prove the desired conclusion. Again, we leave the verification that the two definable operations are mutually inverse to the reader. □

We can now associate a concrete pospace to each set-generated $\bigvee$-semilattice. For a set-generated frame A with generating set g, Proposition 4.20 shows that we can replace the definable collection of classes of frame morphisms from A to P with the class of generating filters of A. The class of points of the space is the class of generating filters of A, that we write $\mathsf{Pt}(A)$ from now on, and the poset of neighbourhood indices is given by g. We can then define a relation between $\mathsf{Pt}(A)$ and g by letting, for α in $\mathsf{Pt}(A)$ and a in g,

$$\alpha \Vdash a =_{\text{def}} a \in \alpha .$$

A series of routine calculations leads to the following result.

Proposition 4.21. *If A is a set-generated frame with generating set g, then $(\mathsf{Pt}(A), \Vdash, g)$ is a concrete pospace.*

The following notion will be relevant when we discuss the connection between examples of set-generated frames and well-known examples of concrete pospaces in Section 4.6. We say that a set-generated frame A with generating set g **has enough points** if for all a in g and all subsets p of g it holds that

$$(\forall \alpha \in \mathsf{Pt}\, A)\big(\alpha \Vdash a \rightarrow (\exists x \in p) \alpha \Vdash x\big) \rightarrow a \leq \bigvee p .$$

4.5 Inductive definition of frames

The goal of this section is to show how inductive definitions can be exploited in CST to define set-presented frames. Given a set s, we know from Theorem 4.14 that any inductive definition Φ on it determines a closure operator c on $\mathsf{Pow}(s)$ and therefore a set-generated $\bigvee$-semilattice, $(\mathsf{Pow}\, s)_c$. Furthermore, the closure operator c is set-presented and therefore so is the $\bigvee$-semilattice $(\mathsf{Pow}\, s)_c$. This, however, still does not give us a frame. To define frames we consider posets rather than sets and inductive definitions on them that determine a set-presentable nucleus j on $\mathsf{Low}(s)$ and hence a frame. This method seems to have been folklore in pointfree topology, where it is associated with the notions known as coverage [46, Section 2.11] or covering system [56, pages 524 – 525], but not to have been combined explicitly with the theory of inductive definitions. The connection between these notions and inductive definitions seems to have been first worked out in the context of type theory in [20]. Working in CST, we can both use the connection between inductive definitions and coverages, and remain close to the original treatment of the method in pointfree topology.

4.5.1 Posites

Let $\mathbb{P}$ be a poset. Recall that we say that a subclass P of $\mathbb{P}$ is a lower class if it holds that $\delta P \subseteq P$, where $\delta P =_{\mathrm{def}} \{x \in \mathbb{P} \mid (\exists y \in P) x \leq y\}$. Define

$$v \leqq u =_{\mathrm{def}} (\forall y \in v)(\exists x \in u) y \leq x$$

for v and u subsets of $\mathbb{P}$. The next definition, but not the terminology, is taken from [56, pages 524 – 525].

Definition 4.22. Let $\mathbb{P}$ be a poset. We say that a function Cov from $\mathbb{P}$ to $\mathsf{Pow}(\mathsf{Pow}\, \mathbb{P})$ is a **coverage** if the following properties

- $(\forall u \in \mathrm{Cov}\, a) u \subseteq \delta\{a\}$
- Cov is stable, i.e. $(\forall x, y \in \mathbb{P}) y \leq x \rightarrow (\forall u \in \mathrm{Cov}\, x)(\exists v \in \mathrm{Cov}\, y) v \leqq u$,

hold. In this case we say that $(\mathbb{P}, \mathrm{Cov})$ is a **posite**. ♢

From now on, we work with a fixed posite $(\mathbb{P}, \mathrm{Cov})$. A particular class of subsets of $\mathbb{P}$ will be of particular interest.

Definition 4.23. We say that a subclass P of $\mathbb{P}$ is an **ideal** of the coverage if it is a lower class and for all a in $\mathbb{P}$ it holds that

$$(\exists u \in \mathrm{Cov}\, a)(u \subseteq P) \rightarrow a \in P\, .$$

We say that an ideal is a **set-ideal** if it is a set. ♢

We write $\mathsf{Idl}(\mathbb{P})$ for the class of set-ideals of the coverage.

4.5.2 Inductive definitions

We now see how posites are related to inductive definitions. Define

$$\Phi =_{\text{def}} \{(\{y\}, x) \mid x, y \in \mathbb{P},\ x \le y\} \cup \{(u, x) \mid x \in \mathbb{P},\ u \in \mathrm{Cov}\, x\}\,,$$

and observe that Φ is an inductive definition on $\mathbb{P}$ in the sense of [7, Chapter 5], i.e. a subset of $\mathsf{Pow}\,\mathbb{P} \times \mathbb{P}$. This is because $\mathbb{P}$ and Cov are sets. We say that a subclass X of $\mathbb{P}$ is **Φ-closed** if

$$(\forall x \in \mathbb{P})(\forall u \in \mathsf{Pow}\,\mathbb{P})\big((u, x) \in \Phi \wedge u \subseteq X \rightarrow x \in X\big)\,.$$

For p in $\mathsf{Pow}\,\mathbb{P}$, $I(\Phi, p)$ is defined as the smallest Φ-closed class containing p. The key aspect of this definition is that, assuming REA, $I(\Phi, p)$ is a set for all subsets p of $\mathbb{P}$. We now show that ideals are exactly the Φ-closed subsets of $\mathbb{P}$. Let us first make more explicit the notion of Φ-closed class. Let p in $\mathsf{Low}(\mathbb{P})$ and define

$$jp =_{\text{def}} I(\Phi, p)\,.$$

By the definition of Φ, we have that $p \subseteq jp$, that jp is a lower class and that

$$(\forall x \in \mathbb{P})(\forall u \in \mathrm{Cov}\, a)(u \subseteq jp \rightarrow x \in p)\,.$$

Furthermore, if P is any lower class that contains p and such that

$$(\forall x \in \mathbb{P})(\forall u \in \mathrm{Cov}\, a)(u \subseteq P \rightarrow x \in P)$$

then $jp \subseteq P$. Observe that p is a set-ideal if and only if $p = jp$ and therefore

$$\mathsf{Idl}(\mathbb{P}) = (\mathsf{Low}\,\mathbb{P})_j\,.$$

The next lemma will be very useful in what follows.

Lemma 4.24 (Induction principle). *Let p a subset of $\mathbb{P}$, and let P be a subclass of $\mathbb{P}$. If P is an ideal and $p \subseteq P$ then $jp \subseteq P$.*

Proof. This is an immediate consequence of the definition of j using Φ. □

The induction principle leads to a proof of Proposition 2.11 of [46] that can be carried over in $\mathbf{CZF}^+$.

Theorem 4.25 (Johnstone's coverage theorem).

(i) *j is a set-presented nucleus on* $\mathsf{Low}(\mathbb{P})$*,*

(ii) $\mathsf{Idl}(\mathbb{P})$ *is a set-presented frame.*

Proof. First of all observe that, once we prove part (i), we will obtain part (ii) as a consequence of the fact that $\mathsf{Idl}(\mathbb{P}) = (\mathsf{Low}\,\mathbb{P})_j$ and of Proposition 4.16. Regarding part (i), we know that j is a set-presented closure operator on $\mathsf{Low}\,\mathbb{P}$ by Theorem 4.14. Hence we just need to show that it preserves meets, i.e. for all p and q lower sets of $\mathbb{P}$ it holds that

$$jp \cap jq \subseteq j(p \cap q)$$

The proof can be obtained as a variation of the one in [46, page 58]. We only sketch its main steps and leave the details to the reader. Define

$$\begin{aligned}
r &=_{\mathrm{def}} \ j(p \cap q)\,, \\
s &=_{\mathrm{def}} \ \{y \in \mathbb{P} \mid (\forall z \in \mathbb{P})(\forall x \in p)\big((z \leq x \wedge z \leq y) \to z \in r\big)\}\,, \\
t &=_{\mathrm{def}} \ \{x \in \mathbb{P} \mid (\forall z \in \mathbb{P})(\forall y \in s)\big((z \leq x \wedge z \leq y) \to x \in r\big)\}\,.
\end{aligned}$$

Observe that $q \subseteq s$ because both p and q are lower sections. We now show $js \subseteq s$ using the induction principle of Lemma 4.24. It therefore suffices to show that s is an ideal. Firstly, observe that s is a lower set. Secondly, let y in $\mathbb{P}$ and let us prove that

$$(\exists v \in \mathrm{Cov}\, y)(v \subseteq s) \to y \in s\,,$$

so let v in $\mathrm{Cov}(y)$ and assume $v \subseteq s$. To prove that $y \in s$, let z in $\mathbb{P}$, x in p and assume $z \leq x$ and $z \leq y$. From $z \leq y$ and $v \in \mathrm{Cov}(y)$, we have that there is $w \in \mathrm{Cov}(z)$ such that $w \leqq v$, because Cov is stable. Now observe that $w \subseteq r$ and $w \in \mathrm{Cov}(z)$. Since r is an ideal we have that z is in r, as required.

Now now have $p \subseteq t$ because q and p are lower sets. We also have $jt \subseteq t$ following the reasoning used to show $js \subseteq s$. Finally observe that $t \cap s \subseteq r$ by the definition of t. We therefore have

$$jp \cap jq \subseteq jt \cap js \subseteq t \cap s \subseteq r = j(p \cap q)\,,$$

as required. □

Remark. We invite the reader to wonder whether the double-negation nucleus defined in Subsection 4.2.2 can be obtained using a posite and a coverage on it. We will answer this problem in Section 5.6.

We now wish to obtain a more explicit description of the set-generated frame $\mathsf{Idl}(\mathbb{P})$. To isolate its generating set, let us introduce some definitions: define a function γ from

$\mathbb{P}$ to $\mathsf{Idl}(\mathbb{P})$ by letting, for a in $\mathbb{P}$

$$\gamma a =_{\text{def}} j\{a\} \tag{4.4}$$

and then define

$$g =_{\text{def}} \{\gamma(x) \mid x \in \mathbb{P}\}\,.$$

The next lemma will be used frequently in what follows, sometimes without explicit mention.

Lemma 4.26. *If p is in $\mathsf{Idl}(\mathbb{P})$ then $p = \bigvee\{\gamma x \mid x \in p\}$, and therefore g is a generating set for $\mathsf{Idl}(\mathbb{P})$.*

Proof. Direct calculations suffice to prove the claim. □

We now consider the points of $\mathsf{Idl}(\mathbb{P})$, introducing the following notion.

Definition 4.27. We say that a subset χ of $\mathbb{P}$ is a **coverage filter** if

- χ is inhabited,
- χ is an upper subset of $\mathbb{P}$,
- χ is stable, i.e. $(\forall x, y \in \mathbb{P})x \in \chi \,\&\, y \in \chi \to (\exists z \in \chi)z \leq x \,\&\, z \leq y\,,$
- χ is closed, i.e. $(\forall x \in \mathbb{P})(\forall u \in \mathrm{Cov}\, x)x \in \chi \leftrightarrow (\exists y \in u)y \in \chi\,,$

hold. ♢

We now show that this definition isolates the subsets of $\mathbb{P}$ that determine points of $\mathsf{Idl}(\mathbb{P})$. In the proof of the next proposition we use the definition fixed in (4.4) and Lemma 4.26.

Proposition 4.28. *Let $(\mathbb{P}, \mathrm{Cov})$ be a posite. There is a bijective correspondence between coverage filters of $\mathbb{P}$ and generating filters of $\mathsf{Idl}(\mathbb{P})$.*

Proof. We exploit Proposition 4.20 and exhibit a bijective correspondence between coverage filters of $\mathbb{P}$ and set-generated completely prime filters of $\mathsf{Idl}(\mathbb{P})$. Let χ be a coverage filter of $\mathbb{P}$, and define

$$F_\chi =_{\text{def}} \{u \in \mathsf{Idl}(\mathbb{P}) \mid (\exists x \in \chi)x \in u\}$$

We claim that F_χ is a set-generated completely prime filter. First, we show that it is inhabited, because χ is so. Let $a \in \chi$, and observe that $a \in \gamma(a)$ and $\gamma(a) \in \mathsf{Idl}(\mathbb{P})$, hence $\gamma(a) \in F_\chi$, as required. To show that F_χ is an upper subclass is simple, and we leave the proof to the reader. Let us now show that F_χ is meet-closed. If $u \in F_\chi$ and $v \in F_\chi$, then there are x, y in χ such that $x \in u$ and $y \in v$, Since χ is stable, there is $z \in \chi$ such that $z \leq x$ and $z \leq y$. We get $z \in u \cap v$ from the fact that u and v are set-ideals and hence lower sets. Hence $u \cap v \in F_\chi$, as required. To show that F_χ is completely prime will require an application of the induction principle of Lemma 4.24. Let p be a subset of $\mathsf{Idl}(\mathbb{P})$. We need to show that

$$\bigvee p \in F_\chi \to (\exists u \in p) u \in F_\chi \, .$$

Let us assume the antecedent of the implication, so let $a \in \chi$ such that $a \in \bigvee p$. Define

$$P =_{\text{def}} \{ x \in \mathbb{P} \mid x \in \chi \to (\exists u \in p) u \in F_\chi \} \, .$$

The proof of the conclusion of the implication will be obtained in two steps: in the first step we prove that

$$j(\bigcup p) \subseteq P \, , \tag{4.5}$$

and in the second step we observe that

$$a \in P \, . \tag{4.6}$$

Once we performed these two steps, we get the desired conclusion from (4.6), the definition of P and the assumption that $a \in \chi$. For the first step, we prove (4.5) using Lemma 4.24. We just need to show that $\bigcup p \subseteq P$, which is easy, and that P is an ideal, which we now do. First we observe that P is a lower class, and then we need to prove that, for y in $\mathbb{P}$

$$(\exists v \in \operatorname{Cov} y)(v \subseteq P) \to y \in P \, .$$

Let $v \in \operatorname{Cov} y$ and assume $v \subseteq P$. To show that $y \in P$, we use the definition of P, so let us assume $y \in \chi$. Since $v \in \operatorname{Cov} y$ and $y \in \chi$, there exists z in v such that $z \in \chi$, because χ is a coverage filter. The conclusion follows by the definition of P, observing that $z \in P$ and $z \in \chi$. For the second step, we prove (4.6) recalling that we assumed $a \in \bigvee p$. We simply observe that

$$a \in \bigvee p = j(\bigcup p) \subseteq P \, ,$$

by the definition of the supremum operation in $\mathsf{Idl}(\mathbb{P})$ and (4.5).

Vice versa, given a set-generated completely prime F define

$$\chi_F =_{\text{def}} \{x \in \mathbb{P} \mid \gamma x \in F \cap g\}\,.$$

We claim that χ_F is a coverage filter. We omit the proofs that χ_F is inhabited and is an upper set, as they are straightforward. We now show that χ_F is stable. Let x, y in χ_F, so that $\gamma x \in F$ and $\gamma y \in F$. Now observe that

$$\gamma x \wedge \gamma y \leq \bigvee\{\gamma z \mid z \leq x\,, z \leq y\}\,,$$

using $\gamma x = j(\delta x), \gamma y = j(\delta y)$ and the fact that the nucleus j preserves meets of lower sets. Hence, we have that there exists z such that $z \leq x$, $z \leq y$ and $\gamma z \in F$, because F is meet-closed, is an upper class and is completely prime. We have therefore found $z \leq x$, $z \leq y$ such that $z \in \chi_F$, i.e. that χ_F is stable, as desired. To show that χ_F is closed, let x in $\mathbb{P}$ and $u \in \operatorname{Cov} x$. We need to show

$$x \in \chi_F \leftrightarrow (\exists y \in u) y \in \chi_F\,.$$

We only prove the 'left-to-right' implication, as the 'right-to-left' is immediate, recalling that F is an upper class. We apply again the induction principle of Lemma 4.24. Define

$$p =_{\text{def}} \bigvee\{\gamma y \mid y \in u\}$$

and observe that p is a set-ideal and that $\{x\} \subseteq p$. Hence, by Lemma 4.24,

$$\gamma x \subseteq p\,.$$

Since F is an upper class, we get that $p \in F$, and hence the desired conclusion because F is completely prime. Direct calculations, some of which use Lemma 4.26, reveal that the two definable operations are mutually inverse. □

We are therefore justified in introducing a slight abuse of language and writing $\mathsf{Pt}(\mathbb{P})$ for the class of coverage filters of $\mathbb{P}$. We use this abuse in the statement of the next corollary.

Corollary 4.29. *The set-presented frame* $\mathsf{Idl}(\mathbb{P})$ *has enough points if and only if for all* a *in* $\mathbb{P}$ *and all* p *subsets of* $\mathbb{P}$

$$(\forall \chi \in \mathsf{Pt}\,\mathbb{P})\big(a \in \chi \to (\exists x \in p) x \in \chi\big) \to a \in jp$$

holds.

Proof. The claim is a direct consequence of the bijection defined in Proposition 4.28. □

4.5.3 Coverage maps

In the following it will also be convenient to have a simpler description of the morphisms whose domain is a frame of the form $\mathsf{Idl}(\mathbb{P})$. The following notion will help us to do so.

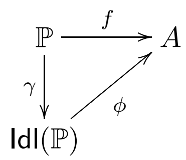
Definition 4.30. Let A be a set generated frame. We say that a function f from $\mathbb{P}$ to A is a **coverage map** if the following properties

- f respects top element, i.e. $\top \leq \bigvee\{f(x) \mid x \in \mathbb{P}\}$,
- f is monotone,
- f respects meets, i.e. $(\forall x, y \in \mathbb{P}) fx \wedge fy \leq \bigvee\{fz \mid z \leq x\,,\ z \leq y\}$,
- f sends covers to joins, i.e. $(\forall x \in \mathbb{P})(\forall u \in \operatorname{Cov} x) fx = \bigvee\{fy \mid y \in u\}$,

hold. ◊

In the following result, γ is the function defined in (4.4) on page 69.

Proposition 4.31. *There is a bijective correspondence between coverage maps from $\mathbb{P}$ to A and frame morphisms from $\mathsf{Idl}(\mathbb{P})$ to A. For every coverage map f from $\mathbb{P}$ to A there exists a unique frame morphism ϕ from $\mathsf{Idl}(\mathbb{P})$ to A such that the following diagram*



commutes, and every ϕ arises in this way.

Proof. We only indicate the definable operations mapping coverage maps into frame morphisms and vice versa, and leave the rest of the verification to the reader. Given a coverage map f from $\mathbb{P}$ to A define a function ϕ_f from $\mathsf{Idl}(\mathbb{P})$ to A by letting, for p in $\mathsf{Idl}(\mathbb{P})$

$$\phi_f(p) =_{\text{def}} \bigvee\{fx \mid x \in p\}$$

The proof that ϕ_f is a frame morphism follows the pattern of the first part of the proof of Proposition 4.28. The claim that ϕ_f makes the appropriate diagram commute follows using Lemma 4.26 and the fact that ϕ_f preserves suprema. Vice versa, given a frame morphism ϕ it is immediate to define a coverage map f_ϕ by letting, for a in $\mathbb{P}$,

$$f_\phi(a) =_{\text{def}} \phi(\gamma a)$$

The proof that f_ϕ is a coverage map now follows the pattern of the second part of the proof of Proposition 4.28. Finally, observe that the two defined operations are mutually inverse. □

4.6 Three spaces

We now present examples of set-generated and set-presented frames. The examples we give are well-known in the literature of pointfree and formal topology [20, 30, 68, 85] and [56, pages 524 – 525] but have not been studied in CST previously. In particular, we aim to transfer to our context some of the results obtained in [30]. When considering frames in CST, we can relate them with important concrete pospaces that can be defined in CST: the Baire space, the Cantor space and, most importantly, the space of Dedekind cuts. One of the reasons for the interest in Dedekind cuts in CST is that their definition differs slightly from the one of Dedekind cuts in intuitionistic set theories of elementary toposes. For example, it is not trivial to show that the Dedekind cuts form a set in **CZF**.

As we will see, for each of these concrete pospaces we will define a set-presented frame such that the considered space is homeomorphic to the space of points of the frame. The method described in Section 4.5 will play an essential role in this process. Our examples will indeed be obtained by first defining posites and then considering the set-presented frames of set-ideals associated with them, as in Theorem 4.25.

4.6.1 Preliminaries

Let s be a set. Define $\mathsf{Seq}(s)$ to be the sets of sequences, i.e. finite lists of elements of s. These can be formally defined as functions with domain a set of the form $\{1, \dots, m\}$ for a natural number m, and codomain s. We will write u_n for the nth element of a sequence u, i.e. for the result of applying a sequence u to n.

We need to introduce some notation associated with sequences. For u in $\mathsf{Seq}(s)$ and a in s, we write $u \cdot a$ for the sequence that is obtained by appending a to u. For a sequence v, we write, $\text{len}(v)$ for its length, which is defined in the standard way. If v is a sequence and n is a natural number such that $n \leq \text{len}(v)$ we define

$$v[n] =_{\text{def}} (v_1, \dots, v_n)\,.$$

This notation allows us to express some notions that will be very useful in the following: for sequences u, v define

$$u \sqsubseteq v =_{\text{def}} (\exists n \in \mathsf{N}) u = v[n]\,. \tag{4.7}$$

We say that u is an **initial segment** of v if $u \sqsubseteq v$. We now consider $\mathsf{Seq}(s)$ as partially ordered by the opposite of the relation defined in (4.7), i.e. we define a partial order on $\mathsf{Seq}(s)$ by letting, for u, v in $\mathsf{Seq}(s)$ $v \leq u =_{\text{def}} u \sqsubseteq v$. In other words we have $v \leq u$ if and only if u is an initial segment of v. We now define a coverage on the poset $\mathsf{Seq}(s)$. For a sequence u define

$$\text{Cov}\, u =_{\text{def}} \{\{u \cdot x \mid x \in s\}\}$$

and observe that this definition determines a coverage on the poset of sequences.

Corollary 4.32. $\mathsf{Idl}(\mathsf{Seq}(s))$ *is a set-presented frame.*

Proof. The claim is a consequence of Theorem 4.25. □

We now isolate the generating set for the frame. For a sequence u define

$$\gamma u =_{\text{def}} \{v \in \mathsf{Seq}(s) \mid u \sqsubseteq v\}$$

and observe that γu is a set-ideal of $\mathsf{Seq}(s)$.

Proposition 4.33. *The set* $\{\gamma u \mid u \in \mathsf{Seq}(s)\}$ *is a generating set for* $\mathsf{Idl}(\mathsf{Seq}(s))$.

Proof. Direct calculation. □

Define $\mathsf{Path}(s)$ as the set of infinite lists of elements of s, i.e. functions from N to s. In the following, we will refer to elements of $\mathsf{Path}(s)$ as paths. Again, we need to introduce some notation that will be used to show how $\mathsf{Path}(s)$ has a structure of concrete pospace. Let f be a path: for n in N define

$$f[n] =_{\text{def}} (f_1, \ldots, f_n)$$

and for a sequence u define

$$u \sqsubset f =_{\text{def}} (\exists n \in \mathsf{N}) u = f[n] \tag{4.8}$$

We say that a sequence u is an **initial segment** of a path f if if $u \sqsubset f$. Note that it could be possible to think of the relation defined in (4.8) as an extension of the one defined in (4.7), since we have that $u \sqsubset f$ when u is an initial segment of f. We obtain a structure of concrete pospace on $\mathsf{Path}(s)$ by considering elements of $\mathsf{Path}(s)$ as points, taking $\mathsf{Seq}(s)$ as the poset of neighbourhoods indices and defining

$$f \Vdash u =_{\text{def}} u \sqsubset f$$

for f in $\mathsf{Path}(s)$ and u in $\mathsf{Seq}(s)$.

Some intuition. The reader may think of spaces like $\mathsf{Path}(s)$ for some set s as a tree growing downwards. Nodes are labeled with natural numbers and elements of the Baire space can be drawn as infinite paths in the tree. For a sequence u in $\mathsf{Seq}(s)$ we have

$$B_u =_{\text{def}} \{f \in \mathsf{Path}\, s \mid u \sqsubset f\},$$

and we refer to B_u as the lower cone determined by u. Neighbourhood indices for the topology can be drawn as the lower cone they determine. Observing that

$$B_{f[n+1]} \subsetneqq B_{f[n]}$$

for all n in N, the reader may then think of the sequence of lower cones determined by the initial segments of f as a progressive approximation to f.

The concrete pospaces given by $\mathsf{Path}(\mathsf{N})$ and $\mathsf{Path}(2)$ are the well-known Baire and Cantor space [84, Section 10.1].

4.6.2 The Baire and Cantor spaces

A pointfree description of the Baire space can be given by considering sequences in N. Define B as the class of set-ideals in $\mathsf{Seq}\,\mathsf{N}$, i.e. define

$$\mathsf{B} =_{\text{def}} \mathsf{Idl}(\mathsf{Seq}\,\mathsf{N}),$$

so that Theorem 4.25 implies that B is a set-presentable frame. We now wish to show that the space of points of B is homeomorphic to the Baire space. In view of Proposition 4.28, we begin by defining a bijection between $\mathsf{Pt}(\mathsf{Seq}\,\mathsf{N})$, i.e. the coverage filters of $\mathsf{Seq}(\mathsf{N})$, and $\mathsf{Path}(\mathsf{N})$. We begin by setting up the bijection. Let f in $\mathsf{Path}(\mathsf{N})$. Define

$$\chi_f =_{\text{def}} \{u \in \mathsf{Seq}\,\mathsf{N} \mid u \sqsubset f\},$$

and observe that the χ_f is a linearly ordered subset of $\mathsf{Seq}(\mathsf{N})$. Let χ be a coverage filter of $\mathsf{Seq}(\mathsf{N})$. Define a relation on $\mathsf{N} \times \mathsf{N}$ by letting

$$f_\chi =_{\text{def}} \{(n, m) \in \mathsf{N} \times \mathsf{N} \mid (\exists u \in \chi) u_n = m)\},$$

and observe that f_χ is a set by Restricted Separation. The next lemmas show that we just defined a bijection between $\mathsf{Pt}(\mathsf{Seq}\,\mathsf{N})$ and $\mathsf{Path}(\mathsf{N})$.

Lemma 4.34.

(i) *If f in* $\mathsf{Path}(\mathsf{N})$ *then χ_f is a coverage filter.*

(ii) *If χ is coverage filter then f_χ is in* $\mathsf{Path}(\mathsf{N})$.

Proof. For (i) observe first of all that χ_f is an upper set. To show that χ is stable, let u, v in $\mathsf{Seq}(\mathsf{N})$ and assume $v \in \chi_f$, $w \in \chi_f$. Then either $v \leq w$ or $w \leq v$ because the set χ_f is linearly ordered. Reasoning by disjunction elimination observe that there exists u such that $u \in \chi_f$ and both $u \leq v$ and $u \leq w$, as required. We now show that χ_f is closed. Let u in $\mathsf{Seq}(\mathsf{N})$ and assume $u \in \chi_f$, and so let n in N such that $f[n] = u$. We have $u \cdot f_{n+1} = f[n+1]$. Hence there exists m in N such that $u \cdot m \in \chi_f$. Recalling the definition of the coverage on $\mathsf{Seq}(\mathsf{N})$, this gives us the desired conclusion.

For (ii) first prove by induction that f is a total relation, i.e.

$$(\forall n \in \mathsf{N})(\exists m)(n, m) \in f_\chi\,.$$

For the base case it is sufficient to recall that χ is inhabited. For inductive step observe

$$u \in \chi \to (\exists m \in \mathsf{N}) u \cdot m \in \chi$$

for all u in $\mathsf{Seq}(\mathsf{N})$ because χ is closed. Hence χ is a total relation. We now show that it is functional. Let n in N. Let m in N, u in $\mathsf{Seq}(\mathsf{N})$ on the one hand, and m' in N, u' in $\mathsf{Seq}(\mathsf{N})$ on the other hand, and assume both $u_n = m$, $u \in \chi$ and $u'_n = m'$, $u' \in \chi$. From $u \in \chi$ and $u' \in \chi$ we have that there exists $v \in \chi$ such that $v \leq u$ and $v \leq u'$ and because χ is stable. We have not only $m = v_n$ because $v \leq u$, but also $m' = v_n$, because $v' \leq u$. Hence $m = m'$, as required. □

Lemma 4.35. *If χ is a coverage map then the following properties of f_χ*

(i) $(\forall n \in \mathsf{N}) f_\chi[n] \in \chi\,,$

(ii) $(\forall u \in \mathsf{Seq}\,\mathsf{N}) u \sqsubset f_\chi \to u \in \chi\,,$

(iii) $(\forall u \in \mathsf{Seq}\,\mathsf{N}) u \in \chi \to u \sqsubset f_\chi\,,$

hold.

Proof. We only sketch the proof, that consists of a series of routine calculations. Part (i) can be proved by induction and part (ii) follows from part (i). For (iii) let u in $\mathsf{Seq}(\mathsf{N})$ and assume $u \in \chi$. Let n be the length of u. We have $f_\chi[n] \in \chi$ by part (i). There exists v such that $v \leq u$ and $v \leq f_\chi[n]$ such that $v \in \chi$ because χ is stable. Now observe that either $f_\chi[n] \leq u$ or $u \leq f_\chi[n]$ and hence $u = f_\chi[n]$, because they have the same length, as required. □

Lemma 4.36.

(i) *If* χ *is a coverage map then* $(\chi)_{f_\chi} = \chi$,

(ii) *If* f *is in* $\mathsf{Path}(\mathsf{N})$ *then* $(f)_{\chi_f} = f$.

Proof. For (i) use parts (ii) and (iii) of Lemma 4.35. For (ii) a direct calculation gives the desired conclusion. □

Proposition 4.37. $\mathsf{Pt}(\mathsf{B})$ *is homeomorphic to* $\mathsf{Path}(\mathsf{N})$.

Proof. Lemma 4.36 gives us a bijection between the two spaces. Routine calculations, whose details are omitted here, show that these functions are continuous. □

We now wish to investigate logical equivalents of the sentence asserting that B has enough points following [30]. The principle of **monotone bar induction** holds if for all subsets p of $\mathsf{Seq}(\mathsf{N})$ and for all elements u of $\mathsf{Seq}(\mathsf{N})$ if

- p is monotonic, i.e. $(\forall v \in \mathsf{Seq}\,\mathsf{N})(\forall n \in \mathsf{N}) v \in p \to v \cdot n \in p$,
- p is inductive, i.e. $(\forall v \in \mathsf{Seq}\,\mathsf{N})\big((\forall n \in \mathsf{N})(v \cdot n \in p)\big) \to v \in p$,
- p is a bar above u, i.e. $(\forall f \in B_u)(\exists v \in p) f \in B_v$.

then u is in p.

Theorem 4.38. *The following are equivalent:*

(i) *the monotone bar induction principle holds,*

(ii) B *has enough points.*

Proof. We use Proposition 4.29 and show that (i) is equivalent to the sentence asserting that for all u in $\mathsf{Seq}(\mathsf{N})$ and p subset of $\mathsf{Seq}(\mathsf{N})$

$$(\forall \chi \in \mathsf{Pt}(\mathsf{Seq}\,\mathsf{N}))\big(u \in \chi \to (\exists v \in p) v \in \chi\big) \to u \in jp \tag{4.9}$$

holds. We first show that (i) implies (4.9), so let u be in $\mathsf{Seq}\,\mathsf{N}$ and p be a subset of $\mathsf{Seq}(\mathsf{N})$ and assume the premiss in (4.9). Note that jp is monotonic and inductive because it is an ideal of the coverage and observe that it is a bar above u using the premiss (4.9), which we assumed, and Lemma 4.36. By an application of the monotone bar induction principle, we get $u \in jp$, as required.

To show that (4.9) implies the principle of monotone bar induction, let p subset of $\mathsf{Seq}(\mathsf{N})$ and u in $\mathsf{Seq}(\mathsf{N})$ and assume that p is monotonic and inductive. Hence p is an ideal and therefore $p = jp$. Now assume that p is a bar above u. We have

$$(\forall \chi \in \mathsf{Pt}(\mathsf{Seq}\,\mathsf{N}))\big(u \in \chi \to (\exists v \in p) v \in \chi\big)$$

by Lemma 4.36. Hence $u \in jp$, by an application of (4.9), and therefore $u \in p$ as required. □

We can proceed in an analogous fashion to describe the Cantor space, but now we consider sequences in the set 2. Define

$$\mathsf{C} =_{\text{def}} \mathsf{Idl}(\mathsf{Seq}\,2)\,,$$

so that, by Theorem 4.25, C is a set-presentable frame.

Proposition 4.39. $\mathsf{Pt}(\mathsf{C})$ *is homeomorphic to* $\mathsf{Path}(2)$.

Proof. The proof is completely analogous to the one of Proposition 4.37. It is actually simpler since induction on N is replaced by reasoning by cases on 2 . □

The appropriate principle to consider for the existence of enough points for $\mathsf{Pt}(\mathsf{C})$ can be introduced as follows. The **fan principle** holds if for all subsets p of $\mathsf{Seq}(2)$, and for all elements u of $\mathsf{Seq}(2)$ if p is a bar above u then there exists m in N such that m is a bound for p, i.e.

$$(\forall f \in B_u)(\exists v \in p)(\text{len}(v) \leq m \wedge f \in B_v)\,.$$

Theorem 4.40. *The following are equivalent:*

(i) *the fan principle holds,*

(ii) C *has enough points.*

Proof. We use again Proposition 4.29 and show that (i) is equivalent to

$$(\forall \chi \in \mathsf{Pt}(\mathsf{Seq}\,2))\,(u \in \chi \to (\exists v \in p)v \in \chi) \to u \in jp \tag{4.10}$$

for all u in $\mathsf{Seq}(2)$ and p subsets of $\mathsf{Seq}(2)$. We first show that (i) implies (4.10). Let u be in $\mathsf{Seq}(2)$ and p be a subset of $\mathsf{Seq}(2)$, and assume the antecedent of (4.10). Hence p is a bar above u by the correspondence between coverage filters and elements of $\mathsf{Path}(2)$, and therefore there exists a bound m for p by (i). Define f and g to be the paths with initial segment u and then constantly equal to $\mathsf{0}$ and $\mathsf{1}$, respectively. Recall that p is a bar above u and that both f and g are in B_u. We have u in jp recalling that jp contains δp and arguing by cases on whether the length of u is greater or smaller than m.

For the reverse implication, we can use the argument used in [24, Section 3.2] to show that the bar induction principle implies the fan principle, as follows. Let u be in

$\mathsf{Seq}(2)$, p be a subset of $\mathsf{Seq}(2)$ and assume that p is a bar above u. Hence

$$(\forall \chi \in \mathsf{Pt}(\mathsf{Seq}\,2))\,(u \in \chi \rightarrow (\exists v \in p) v \in \chi)$$

by the correspondence between coverage points and elements of $\mathsf{Path}\,2$. Therefore $u \in jp$ by (i). Now define

$$R =_{\text{def}} \{(u, m) \in \mathsf{Seq}\,2 \times \mathsf{N} \mid (\forall f \in B_u)(\exists v \in p)(\operatorname{len} v \leq m \wedge f \in B_v)\}$$

and

$$P =_{\text{def}} \{u \in \mathsf{Seq}\,2 \mid (\exists m)(u, m) \in R\}$$

Simple calculations show that P is an ideal and that $p \subseteq P$. Hence $jp \subseteq P$ by Lemma 4.24. We have therefore $u \in P$, and hence the existence of a bound for p, as required. □

4.6.3 The Dedekind reals

Here we can follow [46, Section IV.1], but the concrete pospace of we consider is the set $\mathsf{R_d}$ of Dedekind cuts as defined in CST. Rather than recalling in full the definition of Dedekind cuts from [7, Section 3.6], we will just review some notation associated to them and a characterisation that will be sufficient in what follows.

Let p be a subset of Q and define

$$\begin{aligned} p^{<} &=_{\text{def}} \{x \in \mathsf{Q} \mid (\exists y \in p) x < y\} \\ p^{>} &=_{\text{def}} \{x \in \mathsf{Q} \mid (\exists y \in p) y < x\} \end{aligned}$$

The characterisation of Dedekind cuts we need in the following is contained in the next proposition.

Proposition 4.41 (Aczel). *Let p be a subset of Q. Then p is a cut if and only if it holds that*

- *p is inhabited,*
- *$(\mathsf{Q} \setminus p)$ is inhabited,*
- *$p = p^{<}$,*
- *$(\forall x \in \mathsf{Q})(\forall y \in \mathsf{Q})(x < y \rightarrow x \in p \vee y \notin p)$.*

Proof. See Proposition 3.14 of [7]. □

We write $\mathsf{R_d}$ for the class of cuts and recall that Theorem 3.24 of [7] implies that the class $\mathsf{R_d}$ is a set in **CZF**. Now consider the poset of rational numbers Q with usual order. Define Q^- and Q^+ as the posets obtained by adding a bottom and a top element to the poset Q, respectively. We write these elements as $-\infty$ and $+\infty$ respectively. We define a partial order on $\mathsf{Q}^- \times \mathsf{Q}^+$ by letting, for a, b, c, d in $\mathsf{Q}^- \cup \mathsf{Q}^+$

$$(b, c) \leq (a, d) =_{\mathrm{def}} a \leq b \leq c \leq d\,.$$

We also want to have an order relation between rationals and cuts: for p in $\mathsf{R_d}$ and a, d in Q, define

$$a \leq p \quad =_{\mathrm{def}} \quad a \in p\,, \tag{4.11}$$

$$p \leq d \quad =_{\mathrm{def}} \quad d \in (\mathsf{Q} \setminus p)^{>}\,, \tag{4.12}$$

and extend this definition to a, d in $\mathsf{Q}^- \cup \mathsf{Q}^+$ in the obvious way. We can now consider $(\mathsf{R_d}, \Vdash, \mathsf{Q}^- \times \mathsf{Q}^+)$ as a concrete pospace, where the relation between cuts and neighbourhood intervals is given as follows:

$$p \Vdash (a, d) =_{\mathrm{def}} a \leq p \leq d\,.$$

We now define a set-presented frame whose points correspond to the Dedekind cuts, using once again the method explained in Section 4.5. In order to define a coverage on $\mathsf{Q}^- \times \mathsf{Q}^+$, we need some notation: for a, b, c, d in $\mathsf{Q}^- \cup \mathsf{Q}^+$ define

$$(b, c) \prec (a, d) =_{\mathrm{def}} a < b < c < d$$

and, if $a \leq b < c \leq d$, define

$$(a, c).(b, d) =_{\mathrm{def}} (a, d)\,. \tag{4.13}$$

We then introduce a coverage on $\mathsf{Q}^- \times \mathsf{Q}^+$ as follows: for u in $\mathsf{Q}^- \times \mathsf{Q}^+$ define

$$\mathrm{Cov}(u) =_{\mathrm{def}} \{\{v \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid v \prec u\}\} \cup \{\{v, w\} \mid v.w = u\}\}\,.$$

We invite the reader to draw pictures of the sets in $\mathrm{Cov}(u)$ to gain some intuition into the definition just given. We can now define

$$\mathsf{D} =_{\mathrm{def}} \mathsf{Idl}(\mathsf{Q}^- \times \mathsf{Q}^+)$$

so that, from Theorem 4.25, D is a set-presented frame. In view of Proposition 4.28 to set up a bijective correspondence between points of D and elements of $\mathsf{R_d}$, it suffices to define a bijection between coverage filters of $\mathsf{Q}^- \times \mathsf{Q}^+$ and elements of $\mathsf{R_d}$. Let χ be a coverage point of $\mathsf{Q}^- \times \mathsf{Q}^+$ and define

$$p_\chi =_{\text{def}} \{x \in \mathsf{Q} \mid (x, +\infty) \in \chi\}\,.$$

Now let p be a cut and define

$$\chi_p =_{\text{def}} \{(x, y) \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid x \leq p\,,\ p \leq y\}$$

where we made use of the definitions in (4.11) and (4.12).

Lemma 4.42.

(i) *If χ is a coverage filter then p_χ is a cut.*

(ii) *If p is a cut then χ_p is a coverage filter.*

Proof. For (i) we use Proposition 4.41. Let χ be a coverage filter. First of all, to show that p_χ is inhabited, we use the fact that χ is, and hence assume that $(a, d) \in \chi$. Since χ is an upper set, we get $(a, +\infty) \in \chi$, and therefore $a \in p_\chi$, which shows that a is inhabited. To show that $\mathsf{Q} \setminus p_\chi$ is inhabited, consider again $(a, d) \in \chi$. Using the fact that χ is closed, we obtain there is (b, c) such that $a < b < c < d$ and $(b, c) \in \chi$. It is then possible to show that $(d, +\infty) \notin p_\chi$ using that χ is stable. Let us now show that

$$p_\chi = (p_\chi)^<\,.$$

We only show the 'left-to-right' inclusion since the 'right-to-left' one follows directly by the fact that χ is an upper set. Let $a \in p_\chi$, so that $(a, +\infty) \in \chi$. Using the fact that χ is closed we obtain $(b, c) \in \chi$ such that $(b, c) \prec (a, +\infty)$. We thus have reached that $b \in p_\chi$, because χ is an upper set, and that $a < b$, as required, and so $a \in p_\chi)^<$, as required. Finally, we assume $a < d$ and show either $a \in p_\chi$ or $d \notin p_\chi$, i.e. either $(a, +\infty) \in \chi$ or $(d, +\infty) \notin \chi$. We know $(-\infty, +\infty) \in \chi$, because χ is inhabited and is an upper set, and therefore we have by the definition in (4.13)

$$(-\infty, d).(a, +\infty) \in \chi\,.$$

The conclusion now follows from the fact that χ is closed.

We now illustrate how to prove (ii). To show that χ is inhabited, use the fact that both p and $\mathsf{Q} \setminus p$ are. The proof that χ_p is an upper set is immediate. To show that it is stable, first use decidability of the partial order on the rationals, and then reason by disjunction elimination. Finally to show that χ_p is closed, we distinguish between the two possible ways of 'covering' an element of $\mathsf{Q}^- \times \mathsf{Q}^+$. For the first possibility, we let (a,d) of $\mathsf{Q}^- \times \mathsf{Q}^+$ and prove that

$$(\exists u \in \chi_p) u \prec (a,d) \leftrightarrow (a,d) \in \chi_p$$

We only prove the 'right-to-left' implication, as the 'left-to-right' one follows immediately by the fact that χ_p is upper closed. Assume that $(a,d) \in \chi_p$, i.e. $a \leq p$ and $p \leq d$, then we have $a \in p$ and $d \in (\mathsf{Q} \setminus p)^{>}$ by the definitions fixed in (4.11) and (4.12). We obtain $a \in p^{<}$ and $d \in ((\mathsf{Q} \setminus p)^{>})^{>}$ by Proposition 3.13 of [7]. Hence there are b, c such that $a < b < c < d$ and $b \in p$ and $c \in (\mathsf{Q} \setminus p)^{<}$, and therefore we have found $(b,c) \in \chi_p$ such that $(b,c) \prec (a,d)$, as required. For the second possibility, assume $a \leq b < c \leq d$, and show

$$(a,c) \in \chi_p \vee (b,d) \in \chi_p \leftrightarrow (a,d) \in \chi_p$$

Again, the 'left-to-right' one follows immediately by the fact that χ_p is upper closed. For the 'right-to-left' implication, recall that $b < c$ and hence $b \in p \vee c \notin p$, by the fact that p is a cut and the fourth condition in Proposition 4.41. We now reason by disjunction elimination: if $b \in p$ then $(b,d) \in \chi_p$ and if $c \notin p$ then $(a,c) \in \chi_p$, as required. □

Lemma 4.43.

(i) *If χ is a coverage filter then $(\chi)_{p_\chi} = \chi$.*

(ii) *If p is a cut then $(p)_{\chi_p} = p$.*

Proof. For (i) let χ be a coverage filter Let (a,b) in $\mathsf{Q}^- \times \mathsf{Q}^+$. Then we have

$$\begin{aligned}
(\chi)_{p_\chi} &= \{(x,y) \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid x \leq p_\chi\,,\ p_\chi \leq y\} \\
&= \{(x,y) \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid (x,+\infty) \in \chi,\ (-\infty,y) \in \chi\,,\ x < y\} \\
&= \{(x,y) \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid (x,y) \in \chi\}\,, \\
&= \chi
\end{aligned}$$

as required. For (ii) let p be a cut. Then we obtain

$$p_{\chi_p} = \{x \in \mathsf{Q} \mid (x,+\infty) \in \chi_p\} = \{x \in \mathsf{Q} \mid x \leq p\} = \{x \in \mathsf{Q} \mid x \in p\}\,, = p$$

as required. □

Theorem 4.44. $\mathsf{Pt}(\mathsf{D})$ *and* $\mathsf{R_d}$ *are homeomorphic.*

Proof. The bijection between $\mathsf{Pt}(\mathsf{D})$ and $\mathsf{R_d}$ obtained from Lemma 4.42 and Lemma 4.43 can be easily be shown to be continuous, but we omit the details. □

We end this chapter by introducing the notion of **locally compact** set-generated frame and showing that D is locally compact. Remarkably, the notion of local compactness can be formulated in CST very much as familiar in pointfree topology. This is because we know what it means for an element to be the supremum of a class, as discussed in Section 4.2. A study of local compactness for formal topologies and of the remarkable connections with the property for a formal topology to have a set of formal points can be found in [21, 22]. To discuss local compactness for frames, let us introduce some notation: for two sets p, q we write

$$p \subseteq_\omega q$$

if p is a subset of q and p is the image of a function with domain a set of the form $\{1, \ldots, m\}$ for a natural number m. Let A a set generated frame, with generating set g. For a and b in A define

$$a \ll b =_{\text{def}} (\forall q \in \mathsf{Pow}\, A)\big(b \leq \bigvee q \rightarrow (\exists p \in \mathsf{Pow}\, A)(p \subseteq_\omega q \wedge a \leq \bigvee p)\big)$$

Note that for a, b in A the formula $a \ll b$ is not restricted.

Definition 4.45. We say that a set-generated frame A is **locally compact** if for all a in A

$$a = \bigvee \{x \in A \mid x \ll a\}$$

holds. ♢

Remark. The definition may be understood as requiring that, for all a in A, a is the supremum of the class $\{x \in A \mid x \ll a\}$. Of course, subclasses of A need not have a supremum in general.

The following result reduces the problem of establishing local compactness of a set-generated frame to the inspection of its generating set.

Proposition 4.46 (Aczel). *Let A be a set-generated frame with generating set g. For a, b in A we have*

$$a \ll b \leftrightarrow (\forall q \in \mathsf{Pow}\, g)\big(b \leq \bigvee q \rightarrow (\exists p \in \mathsf{Pow}\, g)(p \subseteq_\omega q \wedge a \leq \bigvee p)\big)$$

and therefore A is locally compact if and only if for all a in g the property

$$a = \bigvee\{x \in g \mid x \ll a\}$$

holds.

Proof. Direct calculations, repeatedly exploiting that A is set-generated, lead to the desired conclusion. □

Let $(\mathbb{P}, \mathrm{Cov})$ be a posite. We now wish to obtain more explicit conditions characterising when the set-presented frame $\mathsf{Idl}(\mathbb{P})$ is locally compact.

Lemma 4.47. *Let a, b in $\mathbb{P}$ we have*

$$\gamma(a) \ll \gamma(b) \leftrightarrow (\forall q \in \mathsf{Pow}\,\mathbb{P})(b \in jq \to (\exists p \in \mathsf{Pow}\,\mathbb{P})(p \subseteq_\omega q \wedge a \in jp))\,.$$

Proof. Direct calculations: first use Proposition 4.46 and then Lemma 4.26. □

Lemma 4.47 suggests the introduction of a slight abuse of language: for a, b in $\mathbb{P}$ we will write $a \ll b$ instead of $\gamma(a) \ll \gamma(b)$. Note that this does not lead to confusion, since $\mathbb{P}$ is not a frame. In the next lemma we use this notation.

Lemma 4.48. $\mathsf{Idl}(\mathbb{P})$ *is locally compact if and only if for all a in $\mathbb{P}$ the property*

$$a \in j\{x \in \mathbb{P} \mid x \ll a\}$$

holds.

Proof. The conclusion can be obtained as a consequence of Proposition 4.46, Lemma 4.47 and Lemma 4.26. □

We can finally prove that D is locally compact. The next two lemmas provide us with the necessary preliminaries.

Lemma 4.49. *For all v in $\mathsf{Q}^- \times \mathsf{Q}^+$ we have $v \in j\{u \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid u \prec v\}$.*

Proof. Let v in $\mathsf{Q}^- \times \mathsf{Q}^+$ and define $p =_{\mathrm{def}} \{u \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid u \prec v\}$. Observe that p is in $\mathrm{Cov}(v)$ and that jp is an ideal. Hence the desired conclusion. □

Lemma 4.50. *Let u, v in $\mathsf{Q}^- \times \mathsf{Q}^+$. If $u \prec v$ then $u \ll v$.*

Proof. Assume $u \prec v$. Let q be a subset of $\mathsf{Q}^- \times \mathsf{Q}^+$ and assume v is in jq. Define

$$P =_{\mathrm{def}} \{w \in \mathsf{Q}^- \times \mathsf{Q}^+ \mid u \prec w \to (\exists p)(p \subseteq_\omega q \wedge v \in jq)\}\,.$$

It is not hard to show that P is an ideal and that $q \subseteq P$. Hence $jq \subseteq P$ by Lemma 4.24. Recalling that we assumed $u \prec v$, the conclusion follows observing that $u \in P$, which is a consequence of the assumption $u \in jq$ and of the inclusion $jq \subseteq P$. □

Proposition 4.51. D *is locally compact.*

Proof. Recalling Lemma 4.48, the desired conclusion follows from Lemma 4.49 and Lemma 4.50. □

Chapter 5

Heyting-valued interpretations for CST

5.1 Heyting-valued interpretations

Heyting-valued interpretations for Intuitionistic Set Theory (IST) were originally obtained in [38]. These interpretations were subsequently developed and found a number of proof theoretical applications in [72, 73]. The aim of this chapter is to define Heyting-valued interpretations for Constructive Set Theory (CST) and indicate how they may lead to proof theoretical applications in CST analogous to the ones existing in IST. We indeed prove first relative consistency and independence results at the end of the chapter.

Heyting-valued interpretations for a set theory can obtained in three steps. Firstly, it is necessary to isolate a notion of Heyting algebra that guarantees a valid interpretation of the logical axioms and allows to define a wide range of examples, so as to allow applications. Secondly, given one such Heyting algebra, one needs to isolate a class of 'Heyting-valued sets' to interpret sets and to define the interpretation of formulas. Finally, the validity of all the axioms of the set theory under consideration has to be proved.

If we wish to obtain reinterpretations, the set theory under consideration needs to support all the steps of the construction by allowing the definitions involved in them. By means of illustration of this issue, let us consider the Heyting-valued reinterpretations of the intuitionistic set theory **IZF** obtained in [38]. When considering this set theory, one is naturally led to focus on complete Heyting algebras that are sets, since the assumption of Power Set and Full Separation guarantees the possibility of defining many examples of such a notion. The other two steps of Heyting-valued interpretations can then be performed in analogy to the ones of Boolean-valued interpretations for Classical Set Theory [11]. Some attention is required, however, in order to avoid the assumption of the law of excluded middle. For example, the definition of the class of 'Heyting-valued sets' that is used to interpret the sets of **IZF** involves using a notion of ordinal that is compatible with the axioms of intuitionistic logic. Finally, the validity of the axioms is a consequence of a remarkable interplay: to prove the validity of the interpretation of an axiom of **IZF**, the same axiom is generally exploited while working informally. An important exception to this interplay is represented by the proof of the validity of Collection, that makes use not only of Collection but also of Full Separation [38].

The study of Heyting-valued interpretations reveals in full the differences between IST, on the one hand, and CST, on the other hand. None of the choices made in the context of IST seems appropriate to obtain reinterpretations for CST. First of all, we have already discussed in Chapter 4 why it is appropriate, when working in CST, to consider set-generated and set-presented frames rather than complete Heyting algebras that are sets. We will therefore consider set-generated frames as the basic notion to develop Heyting-valued interpretations for CST. In view of Proposition 4.10, we will actually limit ourselves to set-generated frames A explicitly defined as

$$A =_{\text{def}} (\mathsf{Low}\, s)_j$$

where s is a poset and j is a nucleus on $\mathsf{Low}(s)$. For the second step, we prefer to avoid the use of ordinals and rather exploit inductive definitions to define a class $V^{(A)}$ that is used to interpret sets. This is because inductive definitions have been widely explored in CST [7, Chapter 5], while the notion of ordinal has not. When it comes to defining the interpretation, we will have to pay particular attention to the distinction between arbitrary and restricted formulas that is peculiar to CST and was not considered in IST.

As we will see, the constructive set theory $\mathbf{CZF}^-$ is sufficient to perform these first two steps and the Strong Collection axiom that is part of $\mathbf{CZF}^-$ plays a crucial role in supporting the definitions involved in them. For the third and final step, i.e. the proofs of validity of the axioms, we will consider $\mathbf{CZF}^-$ first and **CZF** at a later stage. This is because $\mathbf{CZF}^-$ is sufficient to prove the validity of its axioms under the interpretation, and in particular of Strong Collection. It does not seem possible instead to show that, assuming Subset Collection, the interpretation of all the axioms of **CZF** is valid without further assumptions on the set-generated frame. Remarkably, we do not have to look very far for a suitable assumption: if A is set-presented then we obtain the desired result.

We will conclude the chapter with applications of Heyting-valued interpretations. Firstly, we give proofs of a relative consistency and of an independence result concerning the law of restricted excluded middle (**REM**) introduced in Subsection 2.2.3. The independence proof answers a question posed in Section 4.5 and shows that the assertion that the double-negation nucleus is set-presented cannot be proved in **CZF**. Secondly, we show how Heyting-valued interpretations for CST allow to transfer at the generalised predicative context one the aspects of the interplay between pointfree topology and Heyting-valued interpretations. We investigate the relationship between internal objects i.e. elements of $V^{(A)}$ that satisfy the interpretation of a given formula, and external objects, i.e. sets or classes in the set theory that is used to define the interpretation. We show that internal coverage filters of a posite $(\mathbb{P}, \mathrm{Cov})$ correspond to external frame morphisms from $\mathsf{Idl}(\mathbb{P})$ to A.

Remark. Some of the results in this chapter are based on a joint work with Peter Aczel [37]. The content of this chapter differs slightly from the one therein also because some minor problems related to the interpretation of arbitrary formulas have been corrected.

5.2 Preparations

From now on we work informally in $\mathbf{CZF}^-$. Let s be a poset, let j be a nucleus on $\mathsf{Low}(s)$ and define

$$A =_{\mathrm{def}} (\mathsf{Low}\, s)_j \, .$$

Following the remarks after Proposition 4.10, we can assume that j extends to a closure operator on $\mathsf{Pow}\, s$ such that $j(\delta p) = jp$ for all p in $\mathsf{Pow}\, s$. Recall that for a subclass P of s we defined

$$\delta P =_{\mathrm{def}} \{ x \in s \mid (\exists y \in P) x \leq y \} \, ,$$

and we decided to refer to subclasses P of s such that $\delta P \subseteq P$ as lower classes.

5.2.1 Lifting the nucleus

The distinction between arbitrary and restricted formulas requires some attention when it comes to define their Heyting-valued interpretation. Restricted formulas will be interpreted as as elements of $(\mathsf{Low}\, s)_j$, i.e. subsets p of s such that $p = jp$. Inspired by the considerations in Section 2.3, we are led to consider subclasses of s to interpret arbitrary formulas. To do so correctly, we need to extend the nucleus j to a definable operator J on lower classes that coincides with j on lower sets and that it inherits its properties. The next definition provides us with a candidate to be such a definable operator. Let P be a lower subclass of s and define

$$JP =_{\text{def}} \bigcup\{jv \mid v \subseteq P\}\,. \tag{5.1}$$

Some intuition. The reader may compare this definition with the one of open class in a topological space given in Section 4.4.

First of all, we check that J and j coincide on the lower subclasses of s that are sets, as in the next lemma.

Lemma 5.1. *For all p in $\mathsf{Low}(s)$, $Jp = jp$ holds.*

Proof. Direct calculations. ☐

We now wish to show that the operator J inherits all the properties of the nucleus j. Remarkably, the Strong Collection axiom of $\mathbf{CZF}^{-}$ plays an important role in proving this. We begin with a lemma.

Lemma 5.2. *Let P be a lower subclass of s. It holds that*

$$(\forall u \subseteq JP)(\exists v)\big(v \subseteq P \wedge u \subseteq jv\big)\,.$$

Proof. Let u be a subset of JP. For an element x of u and a subset v of P define $\phi =_{\text{def}} x \in jv$. We have

$$(\forall x \in u)\big((\exists v)(v \subseteq P \wedge \phi) \wedge (\forall v)(\forall w)\big((v \subseteq w \subseteq P \wedge \phi) \rightarrow \phi[w/v]\big)\big)$$

by the definition of J and the fact that j is monotone. We can now apply Proposition 2.5 and obtain that there is a set v such that

$$v \subseteq P \wedge (\forall x \in u)\phi\,.$$

Unfolding the definition of ϕ we get the desired conclusion. ☐

Observe that the proof of the previous lemma exploits Lemma 2.5, that is a consequence of Strong Collection. The next proposition uses Lemma 5.2 to show that J inherits all the properties of j.

Proposition 5.3. *Let P and Q be lower subclasses of s. The following properties*

(i) $P \subseteq JP$,

(ii) *if* $P \subseteq Q$ *then* $JP \subseteq JQ$,

(iii) $J(JP) \subseteq JP$,

(iv) $JP \cap JQ \subseteq J(P \cap Q)$,

hold.

Proof. Direct calculations suffice to prove (i), (ii) and (iv). Lemma 5.2 allows us to obtain (iii). □

The definition in (5.1) makes sense also for arbitrary subclasses of s and therefore the operator J extends to a definable operator on them. It is straightforward to see that this operator satisfies the properties in (i), (ii) and (iii) of Proposition 5.3 for arbitrary subclasses of s. In particular, observe that in the proof of Lemma 5.2 we never used the assumption that the considered subclass is a lower one. Furthermore, for a subclass P of s we have

$$JP = J(\delta P) ,$$

as direct calculations suffice to show. In general, however, J will not satisfy property (iv) of Proposition 5.3 for arbitrary subclasses, but only for lower ones.

5.2.2 Lifting operations

Recall that $(\mathsf{Low}\, s)_j$ is a set generated frame whose elements are subsets of s and that a subset p of s is in $(\mathsf{Low}\, s)_j$ if and only if $p = jp$. Observe that in this characterisation we do not need to assume that p is a lower subset, since we assumed that j extends to a closure operator on $\mathsf{Pow}\, s$. Let us review the definitions of the meet, join and Heyting implication operations of $(\mathsf{Low}\, s)_j$. For p and q subsets of s such that $p = jp$ and $q = jq$ we have

$$\left.\begin{array}{rcl} p \wedge q &=& p \cap q \,, \\ p \vee q &=& j(p \cup q) \,, \\ p \rightarrow q &=& \{x \in s \mid x \in p \rightarrow x \in q\} \,. \end{array}\right\} \tag{5.2}$$

We want to lift these operations to act on subclasses of s. For P and Q subclasses of s such that $P = JP$ and $Q = JQ$ define

$$\left.\begin{array}{rcl} P \wedge Q &=_{\mathrm{def}}& P \cap Q \,, \\ P \vee Q &=_{\mathrm{def}}& J(P \cup Q) \,. \\ P \rightarrow Q &=_{\mathrm{def}}& \{x \in s \mid x \in P \rightarrow x \in Q\} \,. \end{array}\right\} \tag{5.3}$$

Observe that if P and Q subclasses of s such that $P = JP$ and $Q = JQ$ and that are sets, then the groups of definitions in (5.2) and in (5.3) are compatible because of Lemma 5.1.

Lemma 5.4. *Let P and Q be subclasses of s such that $P = JP$ and $Q = JQ$. The following hold:*

(i) *$P \wedge Q$ is a subclass of s such that $J(P \wedge Q) = P \wedge Q$. If R is a subclass of s such that $JR = R$ then $R \subseteq P \wedge Q$ if and only if $R \subseteq P$ and $R \subseteq Q$.*

(ii) *$P \vee Q$ is a subclass of s such that $J(P \vee Q) = P \vee Q$. If R is a subclass of s such that $JR = R$ then $P \vee Q \subseteq R$ if and only if $P \subseteq R$ and $Q \subseteq R$.*

(iii) *$P \to Q$ is a subclass of s such that $J(P \to Q) = P \to Q$. If R is a subclass of s such that $JR = R$ then $R \subseteq P \to Q$ if and only if $R \wedge P \subseteq Q$.*

Proof. Direct calculations. □

Recall the definition of supremum and infimum operations in $(\mathsf{Low}\, s)_j$. For a set u such that, for all p in u, $p = jp$ we have

$$\left.\begin{array}{rcl} \bigvee u & = & j\left(\bigcup u\right), \\ \bigwedge u & = & \bigcap u . \end{array}\right\} \tag{5.4}$$

When extending these definitions, we need consider families of subclasses of s rather than sets of subclasses of s. This is necessary in order to interpret correctly unrestricted quantifiers. Classes of subsets of s can indeed be used to interpret only formulas of the form $(\forall x)\theta$ or $(\exists x)\theta$, where θ is a restricted formula. We wish instead to interpret arbitrary formulas of the form $(\forall x)\phi$ or $(\exists x)\phi$, where ϕ need not be restricted. This issue was overlooked in [37].

Let $(P_x)_{x\in U}$ be a family of subclasses of s, such that for all x in U we have $P_x = J(P_x)$, define

$$\left.\begin{array}{rcl} \bigvee_{x\in U} P_x & =_{\mathrm{def}} & J\left(\bigcup_{x\in U} P_x\right), \\ \bigwedge_{x\in U} P_x & =_{\mathrm{def}} & \bigcap_{x\in U} P_x . \end{array}\right\} \tag{5.5}$$

Observe that if U is a set and U_a is a set for all a in U then the class $\{U_x \mid x \in U\}$ is a set by Replacement, and that the groups of definitions in (5.4) and (5.5) are compatible by Lemma 5.1.

Lemma 5.5. *Let $(P_x)_{x\in U}$ be a family of subclasses of s such that for all x in U we have $P_x = J(P_x)$. The following hold:*

(i) *$\bigvee_{x\in U} P_x$ is a subclass of s such that $\bigvee_{x\in U} P_x = J\big(\bigvee_{x\in U} P_x\big)$. If R is a subclass of s such that $R = JR$ then $\bigvee_{x\in U} P_x \subseteq R$ if and only if $P_a \subseteq R$ for all a in U.*

(ii) *$\bigwedge_{x\in U} P_x$ is a subclass of s such that $\bigwedge_{x\in U} P_x = J\big(\bigwedge_{x\in U} P_x\big)$. If R is a subclass of s such that $R = JR$ then $R \subseteq \bigvee_{x\in U} P_x$ if and only if $R \subseteq P_a$ for all a in U.*

Proof. Direct calculations. □

5.3 Definition of the interpretation

In this section we define a class of 'Heyting-valued sets' and then spell out the interpretation of formulas, thus completing the second step necessary to obtain Heyting-valued interpretations. The third and final step, i.e the proofs of validity for the axioms of $\mathbf{CZF}^-$ and **CZF**, will occupy Section 5.4 and Section 5.5.

To define a class $V^{(A)}$ to interpret sets, we use an inductive definition working informally within $\mathbf{CZF}^-$. If f is a function we write $\mathrm{dom}\, f$ and $\mathrm{ran}\, f$ for its domain and range, respectively. We define the class $V^{(A)}$ as the smallest class X such that if f is a function with $\mathrm{dom}\, f \subseteq X$ and $\mathrm{ran}\, f \subseteq A$ then $f \in X$. This inductive definition determines a class within $\mathbf{CZF}^-$ by Theorem 5.1 of [7]. It is worth highlighting the content of this inductive definition as a lemma.

Lemma 5.6. *Let a be a function. If it holds that*

- $\mathrm{dom}\, a \subseteq V^{(A)}$,

- $(\forall x \in \mathrm{dom}\, a)\ ax \in A$,

then a is an element of $V^{(A)}$.

Proof. The statement is a direct consequence of the inductive definition of $V^{(A)}$. □

When defining the Heyting-valued interpretation of a constructive set theory, it is appropriate to distinguish carefully between the object theory, i.e. the theory that is interpreted, and metatheory, i.e. the theory in which the interpretation is defined. The distinction is particularly subtle here because both the object theory and the metatheory are constructive set theories. Our metatheory is the constructive set theory $\mathbf{CZF}^-$. Although we work informally, we keep the notational conventions used elsewhere in this thesis and reserve the letters x, y, z, u, v, w (possibly with indexes or subscripts) for variables.

The object theories we consider are $\mathbf{CZF}^-$ and extensions of it. In order to define the Heyting-valued interpretation, it is convenient to extend the language $\mathcal{L}$ defined in Section 2.1 and assume that the theories we interpret are formulated in such an extension. Define the language $\mathcal{L}^{(A)}$ to be the extension of the language $\mathcal{L}$ with constants $a, b, c, \ldots$ for elements elements $a, b, c, \ldots$ in $V^{(A)}$. Observe that the symbol a plays two roles

- it is a constant of the object language $\mathcal{L}^{(A)}$,

- it denotes a set in $V^{(A)}$ in the metatheory.

Hence we identify an element a in $V^{(A)}$ with the constant a of the language $\mathcal{L}^{(A)}$. It is convenient to assume $\mathcal{L}^{(A)}$ as the object language for the set theories we interpret.

A convention. If ϕ is a formula of $\mathcal{L}^{(A)}$ with $\mathrm{FV}\phi = \{x\}$ then we understand x both as a variable in the object language and as a variable in the metalanguage. This abuse of language could be formally avoided introducing notation to distinguish between the elements of $V^{(A)}$ in the metatheory and constants of the language $\mathcal{L}^{(A)}$ using the so-called 'Quine corner convention' [53, Chapter I, §14].

Let a in $V^{(A)}$. Observe that for all $x \in \mathrm{dom}\, a$, ax is a subset of s such that $j(ax) = ax$. Let $(P_x)_{x \in \mathrm{dom}\, a}$ be a family of subclasses of s such that for all x in a, $J(P_x) = P_x$ and define

$$\begin{aligned} \bigvee_{x:a} P_x &=_{\mathrm{def}} \bigvee_{x \in \mathrm{dom}\, a} ax \wedge P_x \\ \bigwedge_{x:a} P_x &=_{\mathrm{def}} \bigwedge_{x \in \mathrm{dom}\, a} ax \rightarrow P_x \end{aligned}$$

In order to be able to interpret the equality symbol we use a definition by double set recursion, that is allowed in CST [40, Section 2.2]. Here we use double set recursion to define $a =_A b$, for a and b in $V^{(A)}$, so that the following equation

$$a =_A b = \Big(\bigwedge_{x:a} \bigvee_{y:b} x =_A y\Big) \wedge \Big(\bigwedge_{y:b} \bigvee_{x:a} x =_A y\Big) \tag{5.6}$$

holds. We can finally give the definition of the Heyting-valued interpretation of the language $\mathcal{L}^{(A)}$ in Table 5.1.

Proposition 5.7.

(i) *If ϕ is a sentence of $\mathcal{L}^{(A)}$ then $[\![\phi]\!]$ is a subclass of s such that $J[\![\phi]\!] = [\![\phi]\!]$.*

(ii) *If θ is a restricted sentence of $\mathcal{L}^{(A)}$ then $[\![\theta]\!]$ is a subset of s such that $j[\![\theta]\!] = [\![\theta]\!]$, and therefore $[\![\theta]\!]$ is in A.*

Proof. For (i) observe that Proposition 5.4 and Proposition 5.5 give the desired result. For (ii) observe that the operations of the set-generated frame $(\mathsf{Low}\, s)_j$ suffice to define the interpretation of a restricted formula. □

Definition 5.8. We say that a sentence ϕ of $\mathcal{L}^{(A)}$ is **valid** in $V^{(A)}$ if it holds that $[\![\phi]\!] = \top$. We say that a scheme is **valid** if all of its instances with parameters that are elements of $V^{(A)}$ are valid. We say that the Heyting-valued interpretation of a constructive set theory is **valid** if the interpretation of all its axioms and all its axiom schemes is valid. ◊

Table 5.1: Heyting-valued interpretation of the language $\mathcal{L}^{(A)}$.

$$
\begin{aligned}
[\![a = b]\!] &=_{\text{def}} a =_A b\,,\\
[\![\bot]\!] &=_{\text{def}} \bot\,,\\
[\![\phi \wedge \psi]\!] &=_{\text{def}} [\![\phi]\!] \wedge [\![\psi]\!]\,,\\
[\![\phi \vee \psi]\!] &=_{\text{def}} [\![\phi]\!] \vee [\![\psi]\!]\,,\\
[\![\phi \rightarrow \psi]\!] &=_{\text{def}} [\![\phi]\!] \rightarrow [\![\psi]\!]\,,\\
[\![(\forall x \in a)\phi]\!] &=_{\text{def}} \bigwedge_{x:a} [\![\phi]\!]\,,\\
[\![(\exists x \in a)\phi]\!] &=_{\text{def}} \bigvee_{x:a} [\![\phi]\!]\,,\\
[\![(\forall x)\phi]\!] &=_{\text{def}} \bigwedge_{x \in V^{(A)}} [\![\phi]\!]\,,\\
[\![(\exists x)\phi]\!] &=_{\text{def}} \bigvee_{x \in V^{(A)}} [\![\phi]\!]\,.
\end{aligned}
$$

5.4 Towards validity

We continue to work informally in $\mathbf{CZF}^-$, and prove the validity of the structural and set existence axioms. Validity of collection axioms will be considered in Section 5.5. We begin with a simple lemma.

Lemma 5.9. *Let a and b in $V^{(A)}$ and ϕ a formula with* $\mathrm{FV}\phi = \{x\}$. *Then it holds that*

$$[\![\phi[a/x]]\!] \wedge [\![a = b]\!] \leq [\![\phi[b/x]]\!]\,.$$

Proof. By induction on the structure of ϕ. □

Observe that the axioms for the intuitionistic logic and the restricted quantifiers of CST are valid because of Proposition 5.4, Proposition 5.5 and Lemma 5.9.

Proposition 5.10. *Extensionality and Set Induction are valid in $V^{(A)}$.*

Proof. Validity of Extensionality follows by the equivalence in (5.6). Validity of Set Induction is direct consequence of the inductive definition of $V^{(A)}$. □

We now define an embedding from the class of all sets into $V^{(A)}$. This will be used, for example, to prove the validity of Infinity. For a set a, define by set recursion a function $\widehat{a}$ with domain $\{\widehat{x} \mid x \in a\}$ by letting, for x in a

$$\widehat{a}(\widehat{x}) =_{\mathrm{def}} \top \,,$$

and observe that $\widehat{a}$ is in $V^{(A)}$ by Lemma 5.6. This embedding allows us to introduce a notion that will be used in the proof-theoretic applications of Heyting-valued interpretations in Section 5.6.

Definition 5.11. We say that a formula ϕ with $\mathrm{FV}\phi = \{x_1, \ldots, x_n\}$ is **absolute** if for all $a_1, \ldots, a_n$ the equivalence

$$[\![\phi[\widehat{a}_1, \ldots, \widehat{a}_n / x_1, \ldots, x_n]]\!] = \top \;\leftrightarrow\; \phi[a_1, \ldots, a_n / x_1, \ldots, x_n]$$

holds. ♢

Lemma 5.12. *Let a and b be sets. Then it holds that*

$$\begin{aligned} [\![\widehat{a} = \widehat{b}]\!] = \top \quad &\leftrightarrow a = b \,, \\ [\![\widehat{a} \in \widehat{b}]\!] = \top \quad &\leftrightarrow a \in b \,. \end{aligned}$$

Proof. Direct calculations using Set Induction. □

Proposition 5.13. *All restricted formulas are absolute.*

Proof. The proof of Theorem 1.23 in [11] for Boolean-valued interpretations of Classical Set Theory carries over here. In particular, the set generated frame P plays in our context the same role that the complete Boolean algebra 2 plays in the classical context. □

The embedding of the class of all sets in $V^{(A)}$ is useful in the next proposition.

Proposition 5.14. *Pairing, Union, Infinity and Restricted Separation are valid in $V^{(A)}$.*

Proof. The Heyting-valued interpretation of Pairing and Union can be shown to be valid following the proof used in the context of **ZF** or **IZF** [11, 38]. Validity of Infinity follows by embedding an infinite set in $V^{(A)}$. By means of illustration we present the proof of the validity of Restricted Separation in some detail. Let a in $V^{(A)}$ and let θ a restricted formula with $\mathrm{FV}\theta = \{x\}$. Define a function b with the same domain as a by letting, for x in $\mathrm{dom}\, a$,

$$b(x) =_{\mathrm{def}} a(x) \wedge [\![\theta]\!] \,.$$

By part (ii) of Proposition 5.7 and Restricted Separation bx is a set and $j(bx) = b(x)$ for all x in $\mathrm{dom}\, a$. Hence we have that b is in $V^{(A)}$. Let x in $\mathrm{dom}\, a$. Observe that x is in $\mathrm{dom}\, b$ and $ax \wedge [\![\theta]\!] \leq b(x)$. Hence

$$a(x) \wedge [\![\theta]\!] \leq [\![x \in b]\!]$$

and this implies the validity of $(\forall x \in a)(\theta \rightarrow x \in b)$, as direct calculations show. Now let x in $\mathrm{dom}\, b$ and observe that $x \in \mathrm{dom}\, a$ and that $b(x) \leq a(x) \wedge [\![\theta]\!]$ by the definition of b. Hence we obtain

$$b(x) \leq [\![x \in a]\!] \wedge [\![\theta]\!] \,,$$

by the definition of the interpretation. Validity of $(\forall x \in b)(x \in a \wedge \theta)$ follows now by direct calculations and the definition of the Heyting-valued interpretation. □

5.5 Collection axioms

We complete the proofs of validity by considering Strong Collection and Subset Collection. As discussed in Section 5.1, we consider $\mathbf{CZF}^-$ first and **CZF** later.

5.5.1 Strong Collection

Until now, proofs of validity for the axioms of $\mathbf{CZF}^-$ have been quite straightforward, as we essentially transferred the proofs existing in the literature for **ZF** and **IZF**. The only exception has been Restricted Separation, that required some care when defining the interpretation of formulas. The presence of Restricted Separation instead of Full Separation in $\mathbf{CZF}^-$ has effects also when we consider the validity for Strong Collection. This time however, the issue arises in the use of separation axioms to prove the validity of other axioms. In [38] the validity of the Collection axiom of **IZF** was obtained using Full Separation, and therefore that proof is not useful for our goals and does not seem amenable of a simple modification replacing the application of Full Separation with the use of Restricted Separation. Remarkably, we can prove the validity of Strong Collection without assuming Full Separation, but rather exploiting Strong Collection and its consequences proved in Section 2.5. The next two lemmas play a crucial role in obtaining the desired result.

Lemma 5.15. *Let a in $V^{(A)}$ and let ϕ be a formula of $\mathcal{L}^{(A)}$ with* $\mathrm{FV}\phi = \{x\}$.

$$(\forall u \in A)\big(u \leq [\![(\forall x \in a)\phi]\!] \leftrightarrow (\forall x \in \mathrm{dom}\, a) u \wedge ax \leq [\![\phi]\!]\big)\,.$$

Proof. Direct calculations suffice. □

Lemma 5.16. *Let a in $V^{(A)}$ and let ϕ a formula of $\mathcal{L}^{(A)}$ with* $\mathrm{FV}\phi = \{x, y\}$. *Let p in A and define*

$$P =_{\mathrm{def}} \{(x, y, z) \mid x \in \mathrm{dom}\, a\,,\ y \in V^{(A)}\,,\ z \in p \wedge ax \wedge [\![\phi]\!]\}\,.$$

Assume that $p \subseteq [\![(\forall x \in a)(\exists y)\phi]\!]$. Then there exists a subset r of P such that

$$(\forall x \in \mathrm{dom}\, a) p \wedge a(x) \subseteq j\{z \mid (\exists y)(x, y, z) \in r\}\,.$$

Proof. We begin by setting some definitions that will be convenient in the following. Since we will apply Proposition 2.4, we introduce some notation to make that application more evident. Define

$$Q =_{\mathrm{def}} \{(x, z) \mid (\exists y \in V^{(A)})(x, y, z) \in P\}\,,$$

and then, for x in $\mathrm{dom}\, a$, define $Q_x =_{\mathrm{def}} \{z \mid (x, z) \in Q\}$. For x in $\mathrm{dom}\, a$, v in $\mathsf{Pow}\, s$ define $\psi =_{\mathrm{def}} p \wedge ax \subseteq jv$. Starting from the assumption and using the notation just introduced, by Lemma 5.2 we derive that

$$(\forall x \in \mathrm{dom}\, a)\big((\exists v)(v \subseteq Q_x \wedge \psi) \wedge (\forall v)(\forall w)(v \subseteq w \wedge \psi \rightarrow \psi[w/v])\big)\,.$$

We can now apply Proposition 2.4 and obtain a function f with domain $\mathrm{dom}\, a$ such that

$$(\forall x \in \mathrm{dom}\, a)\big(fx \subseteq Q_x \wedge \psi[fx/v]\big)\,. \tag{5.7}$$

We will want to apply Proposition 2.3 and therefore we introduce some more definitions. Define

$$q =_{\mathrm{def}} \{(x, z) \mid x \in \mathrm{dom}\, a\,,\ z \in fx\}$$

and, for x in $\mathrm{dom}\, a$, y in $V^{(A)}$ and z in s define $\xi =_{\mathrm{def}} (x, y, z) \in P$. By the definitions just introduced and (5.7) we obtain

$$(\forall (x, z) \in q)(\exists y)\xi\,.$$

Therefore we can apply Proposition 2.3 and get a function g with domain q such that

$$(\forall (x, z) \in q)\big((\exists y)(y \in g(x, z)) \wedge (\forall y \in g(x, z))\xi\big)\,.$$

Once we define

$$r =_{\text{def}} \{(x, y, z) \mid (x, z) \in q\,,\ y \in g(x, z)\}\,,$$

the desired conclusion is reached with routine calculations, using the fact that j is monotone. □

Proposition 5.17. *Strong Collection is valid.*

Proof. We will use the same notation and definitions used in Lemma 5.16. Let a in $V^{(A)}$ and let ϕ be a formula with $\mathrm{FV}\phi = \{x, y\}$. Let p in A and assume that

$$p \subseteq [\![(\forall x \in a)(\exists y)\phi]\!]\,.$$

By Lemma 5.16 we get a subset r of P such that

$$(\forall x \in \mathrm{dom}\, a) p \wedge ax \subseteq j\{z \mid (\exists y)(x, y, z) \in r\}\,. \tag{5.8}$$

We now define an element b of $V^{(A)}$ that will give us

$$p \subseteq [\![\mathrm{coll}(x \in a, y \in b, \phi)]\!]\,.$$

Recall from Lemma 5.6 that b needs to be a function in order to be an element of $V^{(A)}$. First of all, define the domain of b to be the set t defined as

$$t =_{\text{def}} \{y \mid (\exists x)(\exists z)(x, y, z) \in r\}\,.$$

Now define the function b with domain t by letting, for y in t

$$by =_{\text{def}} j\{z \mid (\exists x)(x, y, z) \in r\}\,.$$

To conclude the proof, observe that b is in $V^{(A)}$ and that, by (5.8), we have

$$p \subseteq [\![\mathrm{coll}(x \in a, y \in b, \phi)]\!]\,,$$

and this leads to the validity of Strong Collection. □

5.5.2 Subset Collection

We keep working informally in $\mathbf{CZF}^-$, but we now assume that the nucleus j is set-presented, and therefore A is set-presented. Let r be a set presentation for j, that is to say a subset of s such that for a in s and p subset of s

$$a \in jp \leftrightarrow (\exists u \in r)(x \in ju \wedge u \subseteq p)\,.$$

Lemma 5.18. *Let a be in s and let P be a subclass of s. We have*

$$a \in JP \leftrightarrow (\exists u \in r)(a \in ju \wedge u \subseteq P)\,.$$

Proof. Direct calculations suffice to prove the claim. □

Define $g =_{\mathrm{def}} \{j\{x\} \mid x \in s\}$ and recall that g is a generating set for A. The next lemma is proved assuming Subset Collection and exploiting Proposition 2.6.

Lemma 5.19. *Let a, b be in $V^{(A)}$ and let ϕ be a formula with $\mathrm{FV}\phi = \{x, y, z\}$. There exists a subset d of $V^{(A)}$ such that for all z in $V^{(A)}$ and for all p in g if*

$$p \subseteq [\![(\forall x \in a)(\exists y \in b)\phi]\!]\,,$$

then there exists e in d such that $p \subseteq [\![\mathrm{coll}(x \in a, y \in e, \phi)]\!]$.

Proof. Let p in g and assume that

$$p \subseteq [\![(\forall x \in a)(\exists y \in b)\phi]\!] \tag{5.9}$$

In the following we will apply Proposition 2.6 twice. In view of those applications it is convenient to define sets a' and b' as follows:

$$\begin{aligned} a' &=_{\mathrm{def}} \{(x, w') \in \mathrm{dom}\, a \times s \mid w' \in p \cap ax\}\,, \\ b' &=_{\mathrm{def}} \mathrm{dom}\, b \times s\,. \end{aligned}$$

The set a' will be used in the second application of Proposition 2.6, while the set b' will be used in the first. Another definition will be helpful before starting the proof. For x in $\mathrm{dom}\, a$, y in $\mathrm{dom}\, b$ and z in $V^{(A)}$ define the class

$$P_{x,y} =_{\mathrm{def}} ax \wedge by \wedge [\![\phi]\!]\,.$$

We begin the proof by letting x' in a'. By the definition of a' we get x in $\mathrm{dom}\, a$ and w' in $p \cap ax$ such that $x' = (x, w')$. We now define the proposition that will be used in our first application of Proposition 2.6. For q in r, w in q and y' in b' define

$$\psi =_{\mathrm{def}} (\exists y \in \mathrm{dom}\, b)\big(y' = (y, w) \wedge w \in q \cap P_{x,y}\big)\,.$$

From (5.9) and Lemma 5.18 we derive that there is q in r such that $w' \in jq$ and

$$(\forall w \in q)(\exists y' \in b')\psi \, .$$

If we apply Proposition 2.6, we obtain a set c', independent of p, x', q and z, such that there is u in c' for which

$$\mathrm{coll}(w \in q, y' \in u, \psi) \tag{5.10}$$

holds. We can now define the proposition used in our second application of Proposition 2.6. For x', x, w', q and u define

$$\xi =_{\mathrm{def}} (\exists q \in r)(\exists x \in \mathrm{dom}\, a)(\exists w \in p \cap a(x) \cap jq)\chi$$

where $\chi =_{\mathrm{def}} x' = (x, w') \wedge \mathrm{coll}(w \in q, y' \in u, \psi)$. Discharging the assumption of x' in a' that we made at the beginning of the proof, we obtain

$$(\forall x' \in a')(\exists u \in c')\xi \, .$$

A second application of Proposition 5.18 gives us a set c, independent of p and z, such that there is v in c for which

$$\mathrm{coll}(x' \in a', u \in v, \xi) \tag{5.11}$$

holds. The set c allows us to define the set whose existence is claimed in the desired conclusion. For v in c define a function f_v with domain $\mathrm{dom}\, b$ by letting, for y in $\mathrm{dom}\, b$

$$f_v(y) =_{\mathrm{def}} j\{w \mid (y, w) \in \bigcup v\} \, .$$

Define $d =_{\mathrm{def}} \{f_v \mid v \in c\}$ and observe that d is a subset of $V^{(A)}$. To conclude the proof let v in c and assume that it satisfies (5.11). Define $e =_{\mathrm{def}} f_v$ so that we have e in d. To conclude the proof we need to show that

$$p \subseteq [\![\mathrm{coll}(x \in a, y \in e, \phi)]\!]$$

holds. We prove the desired claim in two steps. For the first step, let x in $\mathrm{dom}\, a$ and w' in $p \cap ax$. Using first (5.11) and then (5.10) we obtain that there is q in r such that

$$\big(w' \in jq\big) \wedge \big(q \subseteq \bigcup_{y \in \mathrm{dom}\, e} e(y) \cap [\![\phi]\!]\big) \, .$$

We then get $p \subseteq [\![(\forall x \in a)(\exists y \in e)\phi]\!]$ and this concludes the first step. For the second step, let y in $\mathrm{dom}\, e$ and define

$$t =_{\mathrm{def}} p \cap \{w \in s \mid (y, w) \in \bigcup v\}\,.$$

We have

$$\big(p \cap ey \subseteq jt\big) \wedge \big(t \subseteq [\![(\exists x \in a)\phi]\!]\big)\,,$$

using again (5.11) and (5.10). Therefore we get $p \subseteq [\![(\forall y \in e)(\exists x \in a)\phi]\!]$ and this concludes the second step. Putting together the conclusions reached at the end of the two steps we get the desired result. □

The next proposition is proved assuming Subset Collection.

Proposition 5.20. *Subset Collection is valid in $V^{(A)}$.*

Proof. Let a and b in $V^{(A)}$ and let ϕ be a formula with $\mathrm{FV}\phi = \{x, y, z\}$. We can therefore assume to have a set d as in the conclusion of Lemma 5.19. Then define a function c with domain d by letting, for v in d, $cv =_{\mathrm{def}} \top$. Direct calculations lead to the validity of Subset Collection. □

The next theorem is proved working informally within $\mathbf{CZF}^-$.

Theorem 5.21.

(i) *The Heyting-valued interpretation of* $\mathbf{CZF}^-$ *in $V^{(A)}$ is valid.*

(ii) *Assuming Subset Collection, if A is set-presented, then the Heyting-valued interpretation of* **CZF** *in $V^{(A)}$ is valid.*

Proof. For part (i) combine Proposition 5.10, Proposition 5.14 and Proposition 5.17. For (ii) observe that it follows from part (i) and Proposition 5.20 □

5.6 Proof-theoretic applications

We now present first proof-theoretic applications for the Heyting-valued interpretations developed in the previous sections. The results given here arise by considering the set-generated frame P and the double-negation nucleus on it. Observe that the set $\mathsf{1}$ can be seen also as a poset, with discrete partial order. We have that any subset $\mathsf{1}$ is a lower set of the discrete partial order and therefore $\mathsf{Pow}\,\mathsf{1}$ and $\mathsf{Low}\,\mathsf{1}$ are identical. Recall from Subsection 4.2.2 that, for p in P, we defined the double-negation nucleus as

$$jp =_{\mathrm{def}} \{x \in \mathsf{1} \mid \neg\neg x \in p\}\,.$$

We write P_j for the set-generated frame determined by the nucleus j. Recall that p in P is in P_j if and only if

$$p = \{x \in \mathsf{1} \mid \neg\neg x \in p\}\,.$$

We now extend the nucleus j to an operator J on subclasses of $\mathsf{1}$ following the definition in (5.1). Let P be a class contained in $\mathsf{1}$ and define

$$JP =_{\text{def}} \bigcup\{jv \mid v \subseteq P\}\,.$$

Remark. It is important to observe that

$$\{x \in \mathsf{1} \mid \neg\neg x \in P\} \subseteq JP\,,$$

but it does not seem possible to prove the reverse inclusion without further assumptions on P. In other words, J is the double-negation on subsets, but J it is not necessarily the double-negation on subclasses.

5.6.1 A relative consistency result

Let us now consider $V^{(\mathsf{P}_j)}$ and the restricted excluded middle scheme (**REM**) introduced in Section 2.2. Observe that **REM** is equivalent to the sentence

$$(\forall v \in \mathsf{P})(v = \mathsf{1} \vee \neg v = \mathsf{1})$$

In [19] the set theory $\mathbf{CZF}^- + \mathbf{REM}$ was given an interpretation into a semi-classical system **W** that can in turn be interpreted within a pure type theory with W-types. Here we use Heyting-valued interpretations to obtain a more direct interpretation of $\mathbf{CZF}^- + \mathbf{REM}$ into a theory with intuitionistic logic. The next lemma is the key to obtain it.

Lemma 5.22. **REM** *is valid in* $V^{(\mathsf{P}_j)}$.

Proof. Let θ be a restricted sentence. Observe that $\neg\neg(\neg\neg\theta \vee \neg\theta)$ is derivable in intuitionistic logic. For p in P_j define

$$\neg p =_{\text{def}} p \to \bot\,,$$

and observe that

$$\top = \neg\neg(\neg\neg[\![\theta]\!] \cup [\![\neg\theta]\!])\,,$$

by the validity of Heyting-valued interpretations and direct calculations. We have that $[\![\theta]\!]$ is in P_j by Lemma 5.7, and thus $[\![\theta]\!] = \neg\neg[\![\theta]\!]$. We therefore obtain

$$\top = [\![\theta \vee \neg\theta]\!] ,$$

which shows the validity of **REM**. □

Lemma 5.23. *The Heyting-valued interpretation of* $\mathbf{CZF}^- + \mathbf{REM}$ *in* $V^{(\mathsf{P}_j)}$ *is valid.*

Proof. The claim is a consequence of Lemma 5.22 and part (i) of Theorem 5.21. □

To introduce the next definition we recall that, by standard coding, for a set theory **T** there is a sentence $\mathrm{Con}(\mathbf{T_1})$ in the language of first-order arithmetic asserting the consistency of **T**.

Definition 5.24. We say that a set theory $\mathbf{T_1}$ is **reducible** to another set theory $\mathbf{T_2}$ if the sentence

$$\mathrm{Con}(\mathbf{T_2}) \to \mathrm{Con}(\mathbf{T_1})$$

is provable in first-order arithmetic. ♢

Theorem 1.19 of [11] shows that Boolean-valued interpretations give relative consistency proofs for extensions of **ZF**. That theorem carries over also to Heyting-valued interpretations and therefore we obtain the next result.

Theorem 5.25. $\mathbf{CZF}^- + \mathbf{REM}$ *is reducible to* $\mathbf{CZF}^-$*.*

Proof. The claim is a consequence of Lemma 5.23. □

5.6.2 An independence result

The independence result we prove next was suggested to us by Thierry Coquand, and seems to have been first expected in [39]. Let us now consider the theory **CZF**+**REM**. Recall from Section 2.2 that this set theory has at least the proof-theoretic strength of second-order arithmetic and therefore

$$\mathbf{CZF} + \mathbf{REM} \vdash \mathrm{Con}(\mathbf{CZF}) .$$

Theorem 5.26. *The sentence asserting that the double-negation nucleus is set-presented cannot be proved in* **CZF***.*

Proof. Let ϕ be the sentence asserting that the nucleus j is set-presented and assume

$$\mathbf{CZF} \vdash \phi . \tag{5.12}$$

Theorem 5.21 shows that the Heyting-valued interpretation of **CZF** in $V^{(\mathsf{P}_j)}$ is valid. Furthermore we have seen that **REM** is valid. Combining these two facts we obtain that Con(**CZF**) is valid in $V^{(\mathsf{P}_j)}$. Since Con(**CZF**) is an absolute formula, we have

$$\mathbf{CZF} \vdash \mathrm{Con}(\mathbf{CZF})$$

by Proposition 5.13. But this is a contradiction to Gödel's second incompleteness theorem. We have therefore proved that the assumption (5.12) leads to a contradiction, hence the conclusion. □

Hence, the double-negation nucleus cannot be described using posites and inductive definitions. If this was the case, then the nucleus would indeed be set-presented by Theorem 4.25. A similar conclusion in the context of type theory was obtained in [20], but its proof is rather indirect since it makes use of set-theoretic models of pure type theories.

5.7 Further applications

One of the remarkable aspects of the theory of sheaf toposes is the correspondence between internal notions, i.e. notions defined in the internal logic of a topos, and external notions, i.e. notions defined in the informal setting in which sheaf toposes are considered. Let us discuss this informally in the context of intuitionistic set theory. Here one may consider topological spaces as pairs $(X, \mathcal{O}(X))$ where X is a set of points and $\mathcal{O}(X)$ is a frame of open sets, satisfying the usual conditions. For example, we may consider the topological space of Dedekind cuts, $(R, \mathcal{O}(R))$. It is then well-known that if we consider the topos of sheaves over a topological space $(X, \mathcal{O}(X))$, the internal Dedekind reals correspond to external continuous functions from $(X, \mathcal{O}(X))$ to $(R, \mathcal{O}(R))$ [56, Section VI.8].

For localic toposes the issue is more complex. For a frame A, one would wish to obtain a correspondence between the internal Dedekind reals in the topos of sheaves over A and the external frame morphisms from $\mathcal{O}(R)$ to A. Unfortunately, this relies on the assumption that $\mathcal{O}(R)$ is locally compact [31]. An analogous situation happens also for other spaces: the bar induction and the fan theorem principle are used to obtain representations of internal elements of the Baire and Cantor spaces as frame morphisms from $\mathcal{O}(\mathsf{N}^{\mathsf{N}})$ and $\mathcal{O}(2^{\mathsf{N}})$ to A, respectively [31].

In this section we wish to explore how to improve on this situation in the context of Heyting-valued interpretations. The key observation will be that replacing the concrete spaces R_{d}, N^{N} and 2^{N} with their pointfree counterparts allows us to obtain representation of internal points as external frame morphisms without assuming extra principles. In order to make this precise, we need to introduce some notions. From now on, we will work with a fixed set-generated frame A and consider the Heyting-valued interpretations in $V^{(A)}$.

Definition 5.27. Let ϕ be a formula of $\mathcal{L}^{(A)}$ with free variables $x_1, \ldots, x_n$. We say that elements $a_1, \ldots, a_n$ of $V^{(A)}$ **satisfy** ϕ **in** $V^{(A)}$ if

$$[\![\phi[a_1, \ldots, a_n/x_1, \ldots, x_n]]\!] = \top\,.$$

We say that the elements of a definable collection of classes **represent** the elements of $V^{(A)}$ that satisfy ϕ if there is a definable operation assigning to each class P in the collection an element b_P of $V^{(A)}$ such that for all a in $V^{(A)}$ that satisfy ϕ in $V^{(A)}$ there is a unique class P in the collection such that $[\![a = b_P]\!] = \top$. ◊

We now embark on the characterisation of the internal points. Recall from Section 5.4 that there is an embedding assigning an element $\widehat{a}$ of $V^{(A)}$ to any set a. Let θ be the formula of $\mathcal{L}^{(A)}$ with $\mathrm{FV}\theta = \{x\}$ asserting that x is a posite. By the definition of posite, θ is a restricted formula and therefore if x is a posite then $\widehat{x}$ satisfies θ in $V^{(A)}$ by Proposition 5.13. Let x be a posite, and let ϕ be the formula with a free variable y expressing that y is a coverage filter of $\widehat{x}$. We refer to the elements of $V^{(A)}$ that satisfy ϕ as the **internal points** of the posite x in $V^{(A)}$. We aim to prove the following theorem.

Theorem 5.28. *Let* $(\mathbb{P}, \mathrm{Cov})$ *be a posite. Frame morphisms from* $\mathsf{Idl}(\mathbb{P})$ *to* A *represent internal points of* $(\mathbb{P}, \mathrm{Cov})$ *in* $V^{(A)}$*.*

Proof. In view of Proposition 4.31, it is sufficient to show that coverage maps from $\mathbb{P}$ to A represent internal points of $(\mathbb{P}, \mathrm{Cov})$ in $V^{(A)}$. Given a coverage map f from $\mathbb{P}$ to A, define an element χ_f of $V^{(A)}$ as follows: χ_f is a function with domain $\{\widehat{x} \mid x \in \mathbb{P}\}$ defined by letting, for x in $\mathbb{P}$

$$\chi_f(\widehat{x}) =_{\mathrm{def}} f(x)$$

and observe that χ_f is in $V^{(A)}$ because its domain is a subset of $V^{(A)}$ and its range is a subset of A. The proof that χ_f is a coverage filter of $\widehat{\mathbb{P}}$ is a consequence of the assumption that f is a coverage map. Now, let χ be a filter subset of $\widehat{\mathbb{P}}$ in $V^{(A)}$. We need to find a coverage map f from $\mathbb{P}$ to A such that

$$[\![\chi = \chi_f]\!] = \top\,. \tag{5.13}$$

Define f_χ as the function with domain $\mathbb{P}$ defined by letting, for x in $\mathbb{P}$

$$f_\chi(x) =_{\mathrm{def}} [\![\widehat{x} \in \chi]\!]\,,$$

and observe that f_χ is a coverage map because χ is a filter subset of $\widehat{\mathbb{P}}$. The calculations to show this involve applications of Proposition 5.13, but are straightforward. To show $[\![\chi = \chi_{f_\chi}]\!] = \top$ we use the validity of Extensionality in $V^{(A)}$, as follows. Let x in $\mathbb{P}$ and observe that

$$[\![\widehat{x} \in \chi]\!] = f_\chi(x) = [\![\widehat{x} \in \chi_{f_\chi}]\!] \,.$$

This follows by the definitions we fixed and by the definition of the Heyting-valued interpretation. Finally, to show that f_χ is unique among the maps f for which (5.13) holds, observe that for all coverage maps f we have $f_{\chi_f} = f$ by simple calculations, that use again the absoluteness of restricted formulas. □

We can now obtain a representation of the internal points of the spaces we discussed in Section 4.6.

Corollary 5.29. *Let A be a set-generated frame.*

(i) *Frame morphisms from* B *to A represent internal points of* $\mathsf{Seq}(\mathsf{N})$,

(ii) *Frame morphisms from* C *to A represent internal points of* $\mathsf{Seq}(2)$,

(iii) *Frame morphisms from* D *to A represent internal points of* $\mathsf{Q}^- \times \mathsf{Q}^+$.

Proof. The claims are direct consequences of Theorem 5.28. □

Corollary 5.29 represents the first step towards the application of Heyting-valued interpretations to transfer to constructive set theories the relative consistency and independence results obtained for intuitionistic set theories [31, 72, 73]. These seem to carry over from the fully impredicative to the generalised predicative setting. For example, Heyting-valued interpretations developed in this chapter could be applied to prove the independence from **CZF** of various choice principles, like dependent and countable choice, and of principles of intuitionistic analysis, like the monotone bar induction and fan theorem principles [31].

We expect however Heyting-valued interpretations to allow also further applications. Investigations into notions of real numbers in intuitionistic mathematics provide examples of interesting open problems. In [25] it is shown that, alongside the well-known notions of Cauchy and Dedekind reals, there is also another class of real numbers that is of interest for intuitionistic mathematics: the Cauchy completion of the rationals [7, Section 3.6]. It is known that, assuming the principle of countable choice, the three notions are equivalent [25]. Heyting-valued interpretations for IST have been remarkably applied to show that the Dedekind and the Cauchy reals are distinct using interpretations in which the countable choice principle fails [31]. It is however a challenging open problem to establish whether the Cauchy reals and the Cauchy completion of the rationals are distinct. Heyting-valued interpretations for CST seem a natural method to investigate open problems of this kind.

Chapter 6

Sheaf interpretations for CST

6.1 Introduction

In [78] Dana Scott presented sheaf interpretations for IST in a rather direct style, avoiding the use of advanced category-theoretic methods. Unfortunately, the content of those slides was never published, and the closest material that can be found in the literature seems to be Michael Fourman's paper [29]. The main goal of this chapter is to show how some of Scott's ideas and results can be transferred from IST to CST.

One of the reasons for which we decided to consider Scott's approach rather than Fourman's is to make the transition from IST to CST as simple as possible. While Scott developed sheaf interpretations working informally in IST, Fourman obtained them using the internal logic of sheaf toposes. We considered to pursue also this second, more abstract, approach. This would have required, however, a substantial development of category theory and categorical logic in CST before even attempting to develop sheaf interpretations.

Another reason for following Scott is to emphasise how sheaf interpretations can be obtained in two separate steps. The first step is determined by a small category $\mathbb{C}$ and the second step is determined by a Lawvere-Tierney operator j in the category of presheaves over $\mathbb{C}$. These two steps can be seen as generalisations of interpretations that are well-known in mathematical logic. The first one, to which we will refer as **presheaf interpretation**, echoes Kripke's interpretation of intuitionistic logic [24, Section 5.3] and Cohen's forcing technique [53], since posets are particular examples of small categories. The second one, to which we shall refer as **j-translation**, is instead reminiscent of the Gödel-Gentzen translation of classical logic into intuitionistic logic, since double-negation determines a special example of a Lawvere-Tierney operator.

Before performing these two steps, it is however necessary to define a presheaf $V^{(\mathbb{C})}$ to interpret sets and to define the interpretation of the membership relation. Informally speaking, the presheaf $V^{(\mathbb{C})}$ plays the same role for the category of presheaves on $\mathbb{C}$

that the cumulative hierarchy class plays in the category of classes. This brings us to a further reason to prefer Scott's methods over Fourman's. While Scott's definition of $V^{(\mathbb{C})}$ can be recasted in CST as an inductive definition, Fourman's makes essential use of ordinals, that in CST behave even worse than in Topos Theory.

The abstract content of sheaf interpretations will not be completely lost, however. Concepts of Algebraic Set Theory (AST) will allow us to highlight some of the categorical properties that are used in the definition of sheaf interpretations. For example, we will exhibit some formal analogies between the category of classes and the category of presheaves. Furthermore, the abstract formulation of the property of power classes of AST provides us with insight into the nature of the inductive definition that we shall adopt to isolate the presheaf $V^{(\mathbb{C})}$. We indeed expect that, once a completely satisfactory categorical notion of model for CST is formulated, the interpretations described here may find a completely abstract formulation.

Just like for Heyting-valued interpretations, the process of transferring ideas and results from IST to CST is not straightforward. The experience with Heyting-valued interpretations, however, will provide valuable guidance to develop sheaf interpretations for CST. Two aspects will play, again, a particularly important role: the distinction between restricted and arbitrary formulas, and the presence of collection axioms in CST. Recall, for example, that in Chapter 5 we interpreted restricted sentences and arbitrary sentences as lower sets and lower classes, respectively. In this chapter, we will see how restricted and arbitrary sentences will instead determine set-sieves and sieves, that generalise the notion of lower set and lower class to small categories that are not posets. Similarly, we will see how collection axioms both help us setting up the interpretations and are preserved by them. At the end of the chapter, we will discuss the variety of potential applications and some of the reasons leading to consider interpretations of such generality.

6.2 Presheaves

In this section we introduce the category of presheaves. Following the approach of AST, we prefer to define presheaves as functors into the category of classes, rather than into the category of sets. For the definition of the category of classes, see Section 2.4. For general background information on presheaves, we invite the reader to refer to [56, Chapter I].

6.2.1 A small category

Throughout this chapter, we consider a arbitrary, but fixed small category $\mathbb{C}$. We write $\mathbb{C}_0$ for its set of objects and $\mathbb{C}_1$ for its set of maps. In the following we will use letters

$a, b, c, \ldots$ for elements of $\mathbb{C}_0$ and letters $f, g, h, \ldots$ for elements of $\mathbb{C}_1$. For an element f of $\mathbb{C}_1$, we write $\mathsf{dom}\, f$ and $\mathsf{cod}\, f$ for its domain and codomain, respectively. For a in $\mathbb{C}_0$ we write id_a for the identity map on a. If f and g are such that $\mathsf{dom}\, f = \mathsf{cod}\, g$, we write $f \circ g$ for their composition. We draw such a situation in a diagram as follows:

$$a \xleftarrow{\ f\ } b \xleftarrow{\ g\ } c\,,$$

where $a =_{\mathrm{def}} \mathsf{cod}\, f$, $b =_{\mathrm{def}} \mathsf{dom}\, f$, and $c =_{\mathrm{def}} \mathsf{dom}\, g$. For a in $\mathbb{C}_0$ define

$$\mathsf{y}\, a =_{\mathrm{def}} \{ f \in \mathbb{C}_1 \mid \mathsf{cod}\, f = a \}. \tag{6.1}$$

Observe that for all a in $\mathbb{C}_0$, $\mathsf{y}\, a$ is a set by Restricted Separation. For a, b in $\mathbb{C}_0$ define

$$\mathbb{C}(b, a) =_{\mathrm{def}} \{ f \in \mathbb{C}_1 \mid \mathsf{cod}\, f = a\,,\ \mathsf{dom}\, f = b \}\,.$$

Remark. Most of the results obtained in this chapter can be reobtained by replacing the condition that the category $\mathbb{C}$ is small with the condition that for all a in $\mathbb{C}_0$ the class $\mathsf{y}\, a$, as defined in (6.1) is a set. We prefer to assume that $\mathbb{C}$ is small because the condition is a more familiar one, and there is a plethora of examples thereof.

Some intuition. For those readers who are not familiar with categories, one possible way to develop some intuition is to think of elements of $\mathbb{C}_0$ as 'stages of a process', and of elements of $\mathbb{C}_1$ as 'transitions' between stages. Observe that there may be more than one possible transition between two stages. A map f

$$a \xleftarrow{\ f\ } b$$

can then be thought of as a transition from the stage a to the stage b. For a in $\mathbb{C}_0$, we may think of $\mathsf{y}\, a$ as the set of transitions from the stage a.

Examples. Every poset determines a category. If $\mathbb{P}$ is a poset we define a category whose objects are the elements of $\mathbb{P}$ and whose maps are determined by the partial order. There is a map $a \longleftarrow b$ if and only if $b \leq a$, and in that case any two such maps are considered equal.

6.2.2 The category of presheaves

Let us now consider $\mathbb{C}^{op}$, that is the opposite of the category $\mathbb{C}$. We write $\mathbf{Psh}(\mathbb{C})$ for the category whose objects are **presheaves**, i.e. functors from $\mathbb{C}^{op}$ to **Classes** and whose arrows are **presheaf maps**, i.e natural transformations between presheaves. Let us now present the notation that we will use in relation to presheaves and describe more directly presheaves and presheaf maps.

For a presheaf A, we write A_a for the application of A to a in $\mathbb{C}_0$. For a, b in $\mathbb{C}_0$ and f in $\mathbb{C}(b, a)$, we prefer to leave implicit the function from A_a to A_b, to which we sometimes refer as **restriction**, that is determined by f, and simply write s_f for the result of the application of this function to an element s of A_a. Given these conventions, we may spell out the definition of presheaf by saying that a presheaf A assigns

- a class A_a to each $a \in \mathbb{C}_0$,
- an element $s_f \in A_b$ to each $f \in \mathbb{C}(b, a)$ and $s \in A_a$, where $a, b \in \mathbb{C}_0$,

such that the following properties:

- $s_{id_a} = s$, for all $a \in \mathbb{C}_0$ and $s \in A_a$,
- $(s_f)_g = s_{f \circ g}$ for all $a, b, c \in \mathbb{C}_0$, $f \in \mathbb{C}(b, a)$ and $g \in \mathbb{C}(c, b)$,

hold. For a presheaf map F between two presheaves A and B, we write F_a for the function between the classes A_a and B_a, where a is in $\mathbb{C}_0$. Using this convention, a presheaf map assigns a function F_a to each a in $\mathbb{C}_0$ such that for all $a, b \in \mathbb{C}_0$ the following diagram

$$\begin{array}{ccc} A_a & \xrightarrow{F_a} & B_a \\ \downarrow & & \downarrow \\ A_b & \xrightarrow[F_b]{} & B_b \end{array}$$

commutes, where the vertical arrows are the restrictions of the presheaves A and B. We say that F is **small** if for all a in $\mathbb{C}_0$, F_a is a small map between classes in the sense of Section 2.4.

6.2.3 Examples

Constant presheaves. If A is a class, then we may regard it as the presheaf that assigns:

- $A_a =_{\text{def}} A$ to $a \in \mathbb{C}_0$,
- $s_f \in A_b$ to $f \in \mathbb{C}(b, a)$ and $s \in A_a$, that is defined as

$$s_f =_{\text{def}} s \, ,$$

  where $a, b \in \mathbb{C}_0$.

The required properties follow immediately.

Terminal object. The constant presheaf determined by $\mathbf{1}$ is a terminal object in $\mathbf{Psh}(\mathbb{C})$, just as $\mathbf{1}$ is a terminal object in **Classes**.

Representable presheaves. Let a in $\mathbb{C}_0$. The presheaf $\mathsf{Y}(a)$ assigns:

- $\mathsf{Y}(a)_b =_{\text{def}} \mathbb{C}(b, a)$ to $b \in \mathbb{C}_0$,
- $f_g \in \mathsf{Y}(b)$ to $g \in \mathbb{C}(c, b)$ and $f \in \mathsf{Y}(a)_b$, that is defined as

$$f_g =_{\text{def}} f \circ g\,,$$

  where $b, c \in \mathbb{C}_0$.

The desired properties follow immediately by the axioms for a category.

Sieves. Let a in $\mathbb{C}_0$. We say that a subclass P of $\mathsf{y}\,a$ is a **sieve** on a if for all $b, c \in \mathbb{C}_0$ and for all $f \in \mathbb{C}(b, a)$ it holds that

$$f \in P \leftrightarrow (\forall g \in \mathbb{C}(c, b)) f \circ g \in P\,.$$

For a sieve P on a and $f \in \mathbb{C}(b, a)$ and $s \in A_a$, where $a, b \in \mathbb{C}_0$, define

$$P_f =_{\text{def}} \{g \in \mathsf{y}\,b \mid f \circ g \in P\}$$

and observe that P_f is a sieve on b. We say that a sieve on a is a **set-sieve** if it is a set. The presheaf Ω assigns:

- $\Omega_a =_{\text{def}} \{p \mid p \text{ set-sieve on } a\}$ to $a \in \mathbb{C}_0$,
- $p_f \in \Omega_b$ to $f \in \mathbb{C}(b, a)$ and $p \in \Omega_a$, that is defined as

$$p_f =_{\text{def}} \{g \in \mathsf{y}\,b \mid f \circ g \in p\}$$

  where $a, b \in \mathbb{C}_0$.

The required properties are an immediate consequence of the notion of sieve.

A universal presheaf. The presheaf U assigns:

- $U_a =_{\text{def}} \{s \mid s \text{ function with domain } \mathsf{y}\,a\}$ to $a \in \mathbb{C}_0$,
- $s_f \in U_b$ to $f \in \mathbb{C}(b, a)$ and $s \in U_a$, that is defined as the function with domain $\mathsf{y}\,b$ obtained by letting, for $g \in \mathsf{y}\,b$

$$s_f(g) =_{\text{def}} s(f \circ g)\,,$$

where $a, b \in \mathbb{C}_0$.

Again, the required properties follow by direct calculations that we leave to the reader. Proposition 6.1 on page 113 will justify the terminology 'universal presheaf'.

To conclude this section, it is worth highlighting some connections between sieves and lower classes of a poset. Observe that if $\mathbb{C}$ is a poset, then a sieve on an element a of the poset is nothing but a lower class whose elements are all less or equal to a. In the following, we will show a correspondence between sentences of an extension of the language of CST and sieves. Under that correspondence, restricted sentences will correspond to set-sieves. This is indeed completely analogous to what we did in Section 5.3 in relation to Heyting-valued interpretations.

6.3 From classes to presheaves

The goal of this section is exhibit some formal analogies between the categories **Classes** and **Psh**($\mathbb{C}$), so as to arrive to the definition of a 'cumulative hierarchy presheaf' that will be used to define the interpretation of sets.

6.3.1 A universal object

The first property that **Psh**($\mathbb{C}$) inherits from **Classes** is the existence of a universal object, in the sense of Definition 2.2.

Proposition 6.1. *U is a universal object in* **Psh**($\mathbb{C}$).

Proof. Let A be a presheaf. We first define a presheaf map F from A to U, and then show that F is a monomorphism.

For a in $\mathbb{C}_0$ and s in A_a define $F_a(s)$ as the function with domain $\mathsf{y}\, a$ that maps f in $\mathsf{y}\, a$ into s_f. It is not hard to show that this definition determines a presheaf map. Let us now prove that F is a monomorphism, i.e. that for a in $\mathbb{C}_0$, F_a is injective. Let $s, t \in A_a$ and assume $F_a(s) = F_a(t)$. We need to show that $s = t$. This is rather simple: by the very definition of $F_a(s)$ and the properties of presheaves we have that the application of $F_a(s)$ to id_a is s, and similarly that the application of $F_a(t)$ to id_a is t. Since we assumed that $F_a(s)$ and $F_a(t)$ are extensionally equal functions we have $s = t$, as desired. □

6.3.2 Power presheaves

We now want to transfer from the category of classes to the category of presheaves the property expressed by axiom (**P1**) of Section C.2. In **Classes** this axiom was satisfied because we can define the power class operation. This operation assigns the class $\mathsf{Pow}(A)$, whose elements are the subsets of A, to a class A. Observe that, for every class A, $\mathsf{Pow}(A)$ is a subclass of the class of all sets, that is a universal object in **Classes**.

In the following we will be interested in subpresheaves, that we now define, of the universal presheaf. Let A be a presheaf. We say that A' is a **subpresheaf** of A if A' assigns a class A'_a to $a \in \mathbb{C}_0$ such that the following properties

- $A'_a \subseteq A_a$, for all $a \in \mathbb{C}_0$,
- $s_f \in A'_b$, for all $f \in \mathbb{C}(b, a)$ and $s \in A'_a$, where $a, b \in \mathbb{C}_0$,

hold. The next definition isolates the notion of 'subset' in the category $\mathbf{Psh}(\mathbb{C})$.

Definition 6.2. Let A be a presheaf A. We say that $p \in U_a$ is a **presheaf subset of A at stage a** if, for $b \in \mathbb{C}_0$ and $f \in \mathbb{C}(b, a)$, the following properties

- $p(f) \in \mathsf{Pow}(A_b)$,
- p is closed under restrictions, i.e.

$$(\forall y \in p(f))\,(\forall g \in \mathsf{y}\,b)\; y_g \in p_f(g)\,,$$

where $a \in \mathbb{C}_0$, hold. ◊

Finally, we define the power presheaf operation on presheaves. This is defined by assigning

$$\mathsf{Pow}_{\mathbb{C}}(A)_a =_{\text{def}} \{p \in U_a \mid p \text{ presheaf subset of } A \text{ at stage } a\}$$

to a in $\mathbb{C}_0$. Observe that with this definition, $\mathsf{Pow}_{\mathbb{C}}(A)$ is a subpresheaf of U.

Lemma 6.3. *Let a in $\mathbb{C}_0$ and p in $\mathsf{Pow}_{\mathbb{C}}(A)_a$. For all s in A_a it holds that*

$$s \in p(id_a) \leftrightarrow (\forall f \in \mathsf{y}\,a) s_f \in p(f)\,.$$

Proof. The 'right-to-left' implication is obvious, while the 'left-to-right' direction follows from the fact that p is closed under restrictions. □

Recalling the notion of small map between presheaves from Subsection 6.2.2, and that the product of two presheaves is defined as their 'stage-wise' product [56, Chapter I], it follows that the notion of indexed families of small subobjects, as in Definition 2.1, makes sense also in $\mathbf{Psh}(\mathbb{C})$. Let us however rephrase it explicitly for the convenience of the reader.

Definition 6.4. Let A be a presheaf. For a presheaf I, we say that a subobject $R \rightarrowtail I \times A$ is an **I-indexed family of small subobjects of** A if the composite map $R \rightarrowtail I \times A \xrightarrow{\pi_1} I$ is a small presheaf map. ◊

In order to show that (**P1**) holds in $\mathbf{Psh}(\mathbb{C})$ we introduce an abbreviation and, for a presheaf A, write

$$\mathcal{P}_{\mathbb{C}}(A) =_{\text{def}} \mathsf{Pow}_{\mathbb{C}}(A)\,.$$

The axiom (**P1**) requires to exhibit an A-indexed family of small subobjects of A, $\ni_A$, that plays the role of a 'membership' relation. We define $\ni_A$ as the subpresheaf of $\mathcal{P}_{\mathbb{C}}(A) \times A$ that assigns

$$(\ni_A)_a =_{\text{def}} \{(p, s) \in \mathcal{P}_{\mathbb{C}}(A) \times A \mid s \in p(id_a)\}$$

to $a \in \mathbb{C}_0$. Observe that $\ni_A$ is a subpresheaf of $\mathcal{P}_{\mathbb{C}}(A) \times A$ by Lemma 6.3. Furthermore, $\ni_A$ is indeed an A-indexed family of small subobjects of A, because for $a \in \mathbb{C}_0$ and $p \in \mathcal{P}_{\mathbb{C}}(A)_a$ the fiber of $(\ni_A \circ \pi_1)_a$ is a set.

Proposition 6.5. *Let A be a presheaf. The presheaf $\mathcal{P}_{\mathbb{C}}(A)$ and the A-indexed family of subobjects $\ni_A$ are such that for any I-indexed family of small subobjects $R \rightarrowtail I \times A$ there exists a unique presheaf map $I \xrightarrow{F} \mathcal{P}_{\mathbb{C}}(A)$ that makes*

$$\begin{array}{ccc} R & \longrightarrow & \ni_A \\ \downarrow & & \downarrow \\ I \times A & \xrightarrow[F \times \mathrm{Id}_A]{} & \mathcal{P}_{\mathbb{C}}(A) \times A \end{array}$$

a pullback diagram.

Proof. We only indicate the proof, which is really straightforward. For an I-indexed family of subobjects $R \rightarrowtail I \times A$ we define a presheaf map F that assigns, for a in $\mathbb{C}_0$, a function F_a from I_a to $\mathcal{P}_{\mathbb{C}}(A)_a$.

The function F_a needs to map each element i of I into a presheaf subset of A at stage a. Hence $F_a(i)$ needs to be a function with domain $\mathsf{y}\,a$ that satisfies the properties fixed in Definition 6.2. For f in $\mathbb{C}(b, a)$ define

$$F_a(i)(f) =_{\text{def}} \{y \in A_b \mid (i_f, y) \in R\}\,,$$

where $b \in \mathbb{C}_0$ and we assumed, as we always can, that R is a subpresheaf of $I \times A$. Direct calculations now lead to prove that such definition satisfies all the required properties. We leave these verifications to the reader. □

6.3.3 Cumulative hierarchy presheaf

We can now finally define the presheaf $V^{(\mathbb{C})}$ that will be used to interpret sets. First of all, observe that subpresheaves of the universal presheaf U are partially ordered by 'stage-wise' inclusion, i.e. for A and B subpresheaves of U we let

$$A \subseteq B \leftrightarrow (\forall a \in \mathbb{C}_0) A_a \subseteq B_a \,.$$

We would like to 'define' $V^{(\mathbb{C})}$ as the largest subpresheaf X of U such that

$$\mathsf{Pow}_{\mathbb{C}}(X) \subseteq X$$

in complete analogy with the way the cumulative hierarchy is defined in **Classes**. The above 'definition' can be rephrased as an inductive definition of a class $\widehat{V^{(\mathbb{C})}}$ such that

$$\widehat{V^{(\mathbb{C})}} = \Sigma_{a \in \mathbb{C}_0} V_a^{(\mathbb{C})}$$

To introduce the inductive definition of $\widehat{V^{(\mathbb{C})}}$, let $\widehat{U} =_{\mathrm{def}} \Sigma_{a \in \mathbb{C}_0} U_a$. We inductively define $\widehat{V^{(\mathbb{C})}}$ as the smallest subclass Y of $\widehat{U}$ such that

$$\widehat{\mathsf{Pow}}(Y) \subseteq Y$$

where

$$\widehat{\mathsf{Pow}}(Y) =_{\mathrm{def}} \{(a,p) \mid (\forall b \in \mathbb{C}_0)(\forall f \in \mathbb{C}(b,a)\, p(f) \in \mathsf{Pow}(Y_b) \wedge (\forall y \in p(f))(\forall g \in \mathsf{y}\, b)\, y_g \in p_f(g)\}$$

and we used the abbreviation $Y_b = \{s \mid (b,s) \in Y\}$ for $b \in \mathbb{C}_0$. Once $\widehat{V^{(\mathbb{C})}}$ has effectively been defined we may let, for a in $\mathbb{C}_0$,

$$V_a^{(\mathbb{C})} =_{\mathrm{def}} (\widehat{V^{(\mathbb{C})}})_a \,.$$

The next theorem makes explicit the definition of $V^{(\mathbb{C})}$, and will be constantly used in the following to show that explicitly defined sets are in $V^{(\mathbb{C})}$.

Theorem 6.6 (Scott's definition of $V^{(\mathbb{C})}$). *Let a in $\mathbb{C}_0$. Let s be a function with domain $\mathsf{y}\, a$. Then $s \in V_a^{(\mathbb{C})}$ if and only if for all f in $\mathsf{y}\, a$*

(i) $s(f) \subseteq V_b^{(\mathbb{C})}$,

(ii) *if $t \in s(f)$ then $(\forall g \in \mathsf{y}\, b) t_g \in s_f(g)$.*

where $b =_{\mathrm{def}} \mathsf{dom}\, f$.

Proof. Direct consequence of the inductive definition of $V^{(\mathbb{C})}$. □

Corollary 6.7. $V^{(\mathbb{C})}$ *is a subpresheaf of* U.

Proof. Consequence of Theorem 6.6. □

Let us conclude this section by recalling that $V^{(\mathbb{C})}$ was simply defined as in the body of Theorem 6.6 in [78]. The definition of power presheaves and the results in this section are intended to provide some insight into the definition of $V^{(\mathbb{C})}$, and to make sure that this class could be defined inductively in CST. Note that, by Theorem 5.1 of [7] we only need to assume the axioms of $\mathbf{CZF}^-$ to obtain that $V^{(\mathbb{C})}$ is an inductively defined class.

6.4 Presheaf interpretations

Now that we have a presheaf in which we can interpret sets, we wish to define presheaf interpretations. This will be our first step towards sheaf interpretations. This can be done following the clauses of the **Kripke-Joyal** interpretation of logic in an elementary topos [56, Section VI.6]. A remarkable difference arises when interpreting constructive set theories, in that we need to interpret equality and restricted quantifiers over sets. As we will see, the definition of the presheaf $V^{(\mathbb{C})}$ allows us to do so.

Before the definition of the interpretation, let us make some considerations on the syntax. We will essentially follow the approach used to define Heyting-valued interpretations in Section 5.3. Let us consider a in $\mathbb{C}_0$. We define the language $\mathcal{L}^{(a)}$ to be the extension of the language $\mathcal{L}$ with constants for elements of $V_a^{(\mathbb{C})}$. As usual, we do not distinguish between elements of $V_a^{(\mathbb{C})}$ and the constants added to the language $\mathcal{L}$ and use letters $s, t, r, \ldots$ for them. If ϕ is a formula of $\mathcal{L}^{(a)}$ and f is in $\mathsf{y}\,a$ we define ϕ_f to be the formula obtained from ϕ by leaving free variables unchanged and substituting each constant s appearing in ϕ with the constant s_f. We will refer to this process as **restriction**. Observe that if ϕ is a sentence of $\mathcal{L}^{(a)}$ then ϕ_f is a sentence of $\mathcal{L}^{(b)}$, where $b =_{\text{def}} \mathsf{dom}\, f$.

We can now define the presheaf interpretation of formulas. For a in $\mathbb{C}_0$ and ϕ a sentence of $\mathcal{L}^{(a)}$, we define its presheaf interpretation,

$$a \Vdash \phi \,,$$

by structural induction on ϕ according to Table 6.1.

A convention. Let a in $\mathbb{C}_0$. If ϕ is a formula of $\mathcal{L}^{(a)}$ with $\mathrm{FV}\phi = \{x\}$ then we consider x also as a variable in the metalanguage. If we are considering $a \Vdash \phi$, then we assume that x is in $V_a^{(\mathbb{C})}$. Recall that if f is in $\mathsf{y}\,a$, x remains unchanged in ϕ_f. Hence, when considering $b \Vdash \phi_f$, where $b =_{\mathrm{def}} \mathsf{dom}\, f$, we assume that x is in $V_b^{(\mathbb{C})}$.

Table 6.1: Definition of the presheaf interpretation.

$$
\begin{aligned}
a \Vdash s = t &=_{\mathrm{def}} s = t\,,\\
a \Vdash \bot &=_{\mathrm{def}} \bot\\
a \Vdash \phi \wedge \psi &=_{\mathrm{def}} (a \Vdash \phi) \wedge (a \Vdash \psi)\,,\\
a \Vdash \phi \vee \psi &=_{\mathrm{def}} (a \Vdash \phi) \vee (a \Vdash \psi)\,,\\
a \Vdash \phi \to \psi &=_{\mathrm{def}} (\forall f \in \mathsf{y}\, a)(\mathsf{dom}\, f \Vdash \phi_f \to \mathsf{dom}\, f \Vdash \psi_f)\,,\\
a \Vdash (\exists x \in s)\phi &=_{\mathrm{def}} \big(\exists x \in s\,id_a\big)\ a \Vdash \phi\,,\\
a \Vdash (\forall x \in s)\phi &=_{\mathrm{def}} (\forall f \in \mathsf{y}\, a)\ (\forall x \in sf)\ \mathsf{dom}\, f \Vdash \phi_f\,,\\
a \Vdash (\exists x)\phi &=_{\mathrm{def}} \big(\exists x \in V_a^{(\mathbb{C})}\big)\ a \Vdash \phi\,,\\
a \Vdash (\forall x)\phi &=_{\mathrm{def}} \big(\forall f \in \mathsf{y}\, a\big)\big(\forall x \in V_{\mathsf{dom}\, f}^{(\mathbb{C})}\big)\ \mathsf{dom}\, f \Vdash \phi_f\,.
\end{aligned}
$$

We now want to show that sentences of $\mathcal{L}^{(a)}$ correspond to sieves. The next lemma is the key to do so.

Lemma 6.8 (Monotonicity). *Let a in $\mathbb{C}_0$. Let ϕ be a sentence of $\mathcal{L}^{(a)}$. If $a \Vdash \phi$ then for all f in $\mathsf{y}\, a$ it holds that $b \Vdash \phi_f$, where $b =_{\mathrm{def}} \mathsf{dom}\, f$.*

Proof. By structural induction on ϕ. □

The next definition will provide the desired link between sentences and sieves. For a sentence ϕ of $\mathcal{L}^{(a)}$ define

$$
[\![\phi]\!] =_{\mathrm{def}} \{f \in \mathsf{y}\, a \mid \mathsf{dom}\, f \Vdash \phi_f\}\,.
$$

Proposition 6.9. *Let a in $\mathbb{C}_0$.*

(i) *If ϕ is a sentence of $\mathcal{L}^{(a)}$, then $[\![\phi]\!]$ is a sieve on a.*

(ii) *If θ is a restricted sentence of $\mathcal{L}^{(a)}$, then $[\![\theta]\!]$ is a set-sieve on a, and therefore $[\![\theta]\!]$ is in Ω_a.*

Proof. For part (i) Lemma 6.8 gives the desired conclusion. For part (ii) use structural induction on θ, observing that all the clauses in Table 6.1 defining the presheaf interpretation of a restricted formula are themselves restricted. □

Validity of the interpretation can then be defined in a standard way.

Definition 6.10. Let a in $\mathbb{C}_0$. We say that the presheaf interpretation of a sentence ϕ of $\mathcal{L}^{(a)}$ is **valid** if $a \Vdash \phi$. We say that the presheaf interpretation of an axiom scheme of a constructive set theory is **valid** if for all a in $\mathbb{C}_0$, the interpretation of all of the instances of the axiom scheme with parameters in $V_a^{(\mathbb{C})}$ are valid. The presheaf interpretation of a constructive set theory in $V^{(\mathbb{C})}$ is said to be **valid** if the interpretation of all its axioms and axiom schemes is valid. $\Diamond$

6.5 Towards validity

In this section we prove the validity of the presheaf interpretation of the structural and set existence axioms of **CZF**. Proofs concerning collection axioms are instead postponed until Section 6.6. The next proposition takes care of the structural axioms.

Proposition 6.11. *The presheaf interpretation of Extensionality and Set Induction is valid.*

Proof. To illustrate the reasoning used in the proofs concerning presheaf interpretations, we prove the validity of Extensionality. Validity of Set Induction follows directly by the inductive definition of $V^{(\mathbb{C})}$. For Extensionality, let a in $\mathcal{L}^{(a)}$ and let s and t be in $V_a^{(\mathbb{C})}$. Let f in $\mathsf{y}\, a$, define $b =_{\text{def}} \mathsf{dom}\, f$. We assume that

$$b \Vdash (\forall x \in s_f)(\exists y \in t_f) x = y \wedge (\forall y \in t_f)(\exists x \in s_f) x = y \,, \tag{6.2}$$

and we need to show that

$$b \Vdash s_f = t_f \,.$$

This amounts to showing that $s_f = t_f$, by definition of the presheaf interpretation. Since both s_f and t_f are functions with domain $\mathsf{y}\, b$ it suffices to show that for g in $\mathsf{y}\, b$, it holds that $s_f(g) = t_f(g)$. To do this, we use Extensionality and prove mutual inclusions between the two sets. Let x in $s_f(g)$ and observe that $x \in s(f \circ g)$ by definition of the restriction. Hence

$$c \Vdash (\exists y \in t_{f \circ g}) x = y \,,$$

where $c =_{\text{def}} \mathsf{dom}\, g$, by (6.2). Therefore there is y in $t_{f \circ g}(id_c)$ such that $x = y$ by definition of the presheaf interpretation. But $t_{f \circ g}(id_c) = t(f \circ g)$ and therefore we have $s(f \circ g) \subseteq t(f \circ g)$. The reverse inclusion can be proved using the same pattern of reasoning, and therefore we omit its details. $\square$

We now move on to the set existence axioms of **CZF**, which are considered in the next proposition.

Proposition 6.12. *The presheaf interpretation of Pairing, Union, Infinity and Restricted Separation is valid.*

Proof. We show the validity of Pairing and Restricted Separation by means of illustration. Validity of Union can be proved along similar lines, and validity of Infinity follows by the definition of an embedding of all sets into $V^{(\mathbb{C})}$ analogous to the one defined for Heyting-valued interpretations in Chapter 5.

For Pairing, let a in $\mathbb{C}_0$ and let s and t in $V_a^{(\mathbb{C})}$. We define a function r with domain $\mathsf{y}\,a$ by letting, for f in $\mathsf{y}\,a$,

$$r(f) =_{\text{def}} \{s_f, t_f\}\,.$$

First of all, we show that r is an element of $V_a^{(\mathbb{C})}$ using Theorem 6.6. To apply Theorem 6.6, we need to verify two conditions. The first condition is

$$(\forall f \in \mathsf{y}\,a) r(f) \subseteq V^{(\mathbb{C})}_{\text{dom}\, f}\,.$$

This follows by the assumption that s and t are in $V_a^{(\mathbb{C})}$ and by Theorem 6.6. The second condition is

$$(\forall f \in \mathsf{y}\,a)(\forall x \in r(f))(\forall g \in \mathsf{y}(\mathsf{dom}\, f) x_g \in r(f \circ g)\,.$$

Let f in $\mathsf{y}\,a$ and define $b =_{\text{def}} \mathsf{dom}\, f$. Then let x in $r(f)$ and g in $\mathsf{y}\,b$. We need to show that x_g is in $r(f \circ g)$. By the definition of r we have that either $x = s_f$ or $x = t_f$, and therefore x_g is either $s_{f \circ g}$ or $t_{f \circ g}$. In both cases, x_g is in $r(f \circ g)$ by the definition of r and of restrictions in $V^{(\mathbb{C})}$, as required. We now prove that r satisfies the required property stated in the Pairing axiom. This can be done in two parts. The first part of the proof aims at showing that

$$a \Vdash (\forall x)((x = s \vee x = t) \rightarrow x \in r)\,.$$

Let f in $\mathsf{y}\,a$ and define $b =_{\text{def}} \mathsf{dom}\, f$. Now let x in $V_b^{(\mathbb{C})}$ and assume that either $x = s_f$ or $x = t_f$. In both cases we can show that

$$b \Vdash (\exists y \in r_f) x = y\,,$$

because both s_f and t_f are in $r_f(id_c)$, since $r_f(id_c) = r(f)$. The second part of the proof consists in proving that

$$a \Vdash (\forall x \in r)(x = s_f \vee x = t_f)\,.$$

Again, let f in $\mathsf{y}\,a$ and define $b =_{\text{def}} \mathsf{dom}\, f$. For x in $r_f(id_b)$, we have either $x = s_f$ or $x = t_f$ by definition of $r(f)$. Therefore we obtain the desired conclusion, by the definition of the presheaf interpretation. We have thus completed the proof of the validity of Pairing.

For Restricted Separation, let a in $\mathbb{C}_0$, let s in $V_a^{(\mathbb{C})}$ and let θ be a restricted formula of $\mathcal{L}^{(a)}$ with $\mathrm{FV}\theta = x$. Define a function t with domain $\mathsf{y}\,a$ by letting, for f in $\mathsf{y}\,a$

$$t(f) =_{\text{def}} \{x \in s(f) \mid f \in [\![\theta]\!]\} ,$$

and observe that for all f in $\mathsf{y}\,a$, $t(f)$ is a set by Lemma 6.9 and Restricted Separation. We now want to show that t is in $V_a^{(\mathbb{C})}$ using Theorem 6.6. The first requirement, i.e.

$$(\forall f \in \mathsf{y}\,a) t(f) \subseteq V^{(\mathbb{C})}_{\mathsf{dom}\, f} ,$$

follows immediately by the definition of t. For the second requirement, i.e.

$$(\forall f \in \mathsf{y}\,a)(\forall x \in t(f))(\forall g \in \mathsf{y}(\mathsf{dom}\, f) x_g \in t(f \circ g) ,$$

let f in $\mathsf{y}\,a$, let x in t_f, define $b =_{\text{def}} \mathsf{dom}\, f$ and let g in $\mathsf{y}\,b$. We claim that x_g is in $s(f \circ g)$ and that $f \circ g \in [\![\theta]\!]$. These two claims clearly imply that $x_g \in t(f \circ g)$. The first claim follows by the assumption that s is in $V_a^{(\mathbb{C})}$ and Theorem 6.6. The second claim follows instead by the assumption that x is in $t(f)$ and by Lemma 6.8. The proof that t satisfies the property required in the Restricted Separation axiom can again be divided into two parts. The first part aims at showing that

$$a \Vdash (\forall x \in t)(x \in s \wedge \theta) .$$

Let f in $\mathsf{y}\,a$, define $b =_{\text{def}} \mathsf{dom}\, f$ and let x in $t(f)$. Then we have that $x \in s_f(id_b)$ and that $b \Vdash \theta_f$ by the assumption that x is in $t(f)$ and the definition of t. But this shows exactly the desired conclusion, by definition of the presheaf interpretation. For the second part, we need to show that

$$a \Vdash (\forall x \in s)(\theta \rightarrow x \in t) .$$

Once again, let f in $\mathsf{y}\,a$, define $b =_{\text{def}} \mathsf{dom}\, f$, let x in $s(f)$ and assume $b \Vdash \theta_f$. Then we have that f is in $[\![\theta]\!]$ by the very definition of $[\![\theta]\!]$, and thus x is in $t(f)$, as required. This concludes the proof of the validity of Restricted Separation. □

6.6 Collection axioms

Proofs become slightly more complicated when deal with collection axioms, but in the end we are going to prove that the presheaf interpretation of **CZF** is valid.

6.6.1 Strong collection

The next lemma can be seen as a counterpart of Lemma 5.16, that was used to prove the validity of the Heyting-valued interpretation of Strong Collection in Section 5.5.

Lemma 6.13. *Let a be in $\mathbb{C}_0$, let s be in $V_a^{(\mathbb{C})}$, let ϕ be a formula of $\mathcal{L}^{(a)}$ with* $\mathrm{FV}\phi = \{x, y\}$ *and let f be in $\mathsf{y}\, a$. Define $b =_{\mathrm{def}} \mathsf{dom}\, f$ and*

$$Q =_{\mathrm{def}} \{(g, x, y) \mid x \in s(f \circ g)\,,\ y \in V^{(\mathbb{C})}_{\mathsf{dom}\, g}\,,\ \mathsf{dom}\, g \Vdash \phi_{f \circ g}\}$$

If $b \Vdash (\forall x \in s_f)(\exists y)\phi_f$ then there exists a subset q of Q such that

$$(\forall g \in \mathsf{y}\, b)(\forall x \in s(f \circ g))(\exists y)(g, x, y) \in q\,.$$

Proof. We will want to apply Proposition 2.3 and Proposition 2.4, so let us define some abbreviations to make these applications more evident. We first define a formula that will be used to apply Proposition 2.3. For g, x and y define

$$\xi =_{\mathrm{def}} (g, x, y) \in Q\,.$$

We now define a family of classes that will be used when we apply Proposition 2.4. For $g \in \mathsf{y}\, b$ define

$$Q_g =_{\mathrm{def}} \{(x, y) \mid (g, x, y) \in Q\}\,. \tag{6.3}$$

We now begin the actual proof. Assume

$$b \Vdash (\forall x \in s_f)(\exists y)\phi_f\,,$$

as in the hypothesis of the statement we want to prove, and let g in $\mathsf{y}\, b$. The definition of the presheaf interpretation leads us to obtain

$$(\forall x \in s(f \circ g))(\exists y)\xi\,.$$

We can now apply Proposition 2.3 and get a function l with domain $s(f \circ g)$ such that

$$(\forall x \in s(f \circ g))(\exists y)(y \in lx) \wedge (\forall y \in lx)\xi\,.$$

Once we define

$$p =_{\mathrm{def}} \{(x,y) \mid x \in s(f \circ g)\,,\ y \in lx\}\,,$$

we may observe that, by the definition of Q_g in (6.3), we have

$$p \subseteq Q_g \wedge (\forall x \in s(f \circ g))(\exists y)(x,y) \in p\,.$$

We are approaching the application of Proposition 2.4. For a set u, let us define

$$\psi =_{\mathrm{def}} (\forall x \in s(f \circ g))(\exists y)(x,y) \in u\,.$$

By discharging the assumption of l and universally quantifying over g in $\mathsf{y}\,b$, we get

$$(\forall g \in \mathsf{y}\,b)\big(((\exists u)u \subseteq Q_g \wedge \psi) \wedge ((\forall u)(\forall v)(u \subseteq v \subseteq Q_g \wedge \psi) \rightarrow \psi[v/u])\big)\,.$$

We are under the hypothesis of Proposition 2.4, as planned, and therefore we can apply it to get a function m with domain $\mathsf{y}\,b$ such that

$$(\forall g \in \mathsf{y}\,b)(mg \subseteq Q_g \wedge \psi[mg/u])\,.$$

Once we define

$$q =_{\mathrm{def}} \{(g,x,y) \mid g \in \mathsf{y}\,b\,,\ x \in s(f \circ g)\,,\ (x,y) \in mg\}\,,$$

the desired conclusion follows by direct calculations. □

Proposition 6.14. *The presheaf interpretation of Strong Collection is valid.*

Proof. Let a be in $\mathbb{C}_0$, s be in $V_a^{(\mathbb{C})}$ and ϕ be a formula of $\mathcal{L}^{(a)}$ with $\mathrm{FV}\phi = \{x,y\}$ and f be in $\mathsf{y}\,a$. We use the same definitions of Lemma 6.13. Assume that

$$b \Vdash (\forall x \in s(f \circ g))(\exists y)\phi_{f \circ g}$$

and let q be a set that satisfies the conclusion of Lemma 6.13. To obtain an element of $V^{(\mathbb{C})}$, we need to 'close' q under restrictions, so let us define

$$q' =_{\mathrm{def}} \{(g \circ h, x_h, y_h) \mid (g,x,y) \in q\,,\ h \in \mathsf{y}(\mathsf{dom}\,g)\}\,.$$

We now define a function with domain $\mathsf{y}\,b$ by letting, for g' in $\mathsf{y}\,b$

$$t(g') =_{\mathrm{def}} \{y' \mid (\exists x')(g',x',y') \in q'\}$$

Theorem 6.6 and Lemma 6.13 imply then that t is in $V_b^{(\mathbb{C})}$, as wanted. We now show

$$b \Vdash \mathrm{coll}(x \in s_f, y \in t, \phi_f)\,. \tag{6.4}$$

Let g in $\mathsf{y}\,b$ and x in $s(f \circ g)$. Then there is y such that (g, x, y) is in q, again because we assumed that q satisfies the conclusion of Lemma 6.13. We therefore have that y is in $t(g)$ because (g, x, y) is also in q'. Since q is a subset of Q, we have

$$\mathsf{dom}\, g \Vdash \phi_{f \circ g}\,,$$

and this completes a first part of the proof of (6.4). Let g' in $\mathsf{y}\,b$ and y' in $t(g')$. We have that there is x' such that (g', x', y') is in q' by definition of t. So there are (g, x, y) in q and h in $\mathsf{y}(\mathsf{dom}\, g)$ such that

$$(g', x', y') = (g \circ h, x_h, y_h) \tag{6.5}$$

by definition of q'. We have $\mathsf{dom}\, g \Vdash \phi_{f \circ g}$ by the conclusion of Lemma 6.13 and hence

$$\mathsf{dom}\, h \Vdash \phi_{f \circ g \circ h}[x_h, y_h / x, y]$$

by Lemma 6.8. Finally observe that by (6.5)

$$\mathsf{dom}\, g' \Vdash \phi_{f \circ g'}[x', y' / x, y]$$

which completes the second part of the proof of (6.4). □

6.6.2 Subset Collection

The next lemma is analogous to Lemma 5.19, that implied the validity of the Heyting-valued interpretation of Subset Collection. Note, however, that we do not need to introduce any condition asserting that the category is 'set-presented', since we already assumed that it is small.

Lemma 6.15. *Let a in $\mathbb{C}_0$, let s and t be in $V_a^{(\mathbb{C})}$, let ϕ be a formula of $\mathcal{L}^{(a)}$ with* $\mathrm{FV}\phi = \{x, y, z\}$. *There is a set q of functions with domain $\mathsf{y}\,a$ such that for all u in q, and for all f in $\mathsf{y}\,a$, uf is in $V_b^{(\mathbb{C})}$, where $b =_{\mathrm{def}} \mathsf{dom}\, f$. Furthermore, for all f in $\mathsf{y}\,a$ and all z in $V_b^{(\mathbb{C})}$ if*

$$b \Vdash (\forall x \in s_f)(\exists y \in t_f)\phi_f$$

then there exists u in q such that

$$b \Vdash \mathrm{coll}(x \in s_f, y \in uf, \phi)$$

where $b =_{\mathrm{def}} \mathsf{dom}\, f$.

Proof. We will want to apply Proposition 2.6. We therefore set up the appropriate definitions of sets and formulas to make the application simpler. For f in $\mathsf{y}\, a$, define

$$m^f =_{\mathrm{def}} \{(g, x) \mid g \in \mathsf{y}\, b\,,\ x \in s(f \circ g)\}$$

where $b =_{\mathrm{def}} \mathsf{dom}\, f$. Then define

$$n =_{\mathrm{def}} \{(f, g, y) \mid f \in \mathsf{y}\, a\,,\ g \in \mathsf{y}\, b\,,\ y \in t(s \circ g)\}\,.$$

At various points in the proof we will have to 'close' sets under restrictions in order to obtain elements of $V^{(\mathbb{C})}$. The presence of a number of equalities in the formulas in (6.6) and (6.7) below will be useful when we will deal with sets that have been obtained by 'closing' under restrictions. For f in $\mathsf{y}\, a$, g' in $\mathsf{y}(\mathsf{dom}\, f)$, x' in $s(f \circ g')$ and y'' define

$$\xi' =_{\mathrm{def}} c \Vdash \phi_{f \circ g'}[x', y'', z_{g'}/x, y, z]$$

and for v in n, f'' in $\mathsf{y}\, a$, g'' in $\mathsf{y}(\mathsf{dom}\, f'')$, y'' define

$$\xi'' =_{\mathrm{def}} v = (f'', g'', y'') \wedge (f = f'') \wedge (g' = g'') \tag{6.6}$$

and finally $\xi =_{\mathrm{def}} \xi' \wedge \xi''$. For u in m^f, define

$$\psi' =_{\mathrm{def}} (\exists g' \in \mathsf{y}\, b)(\exists x' \in s(f \circ g'))\ u = (g', x') \tag{6.7}$$

and, for v in n, define

$$\psi'' =_{\mathrm{def}} (\exists f'' \in \mathsf{y}\, a)(\exists g'' \in \mathsf{y}\, b)(\exists y'' \in t(f'' \circ g''))\ \xi\,.$$

Finally, define $\psi =_{\mathrm{def}} \psi' \wedge \psi''$. We can now start the proof. Let f in $\mathsf{y}\, a$, define $b =_{\mathrm{def}} \mathsf{dom}\, f$, let z in $V_b^{(\mathbb{C})}$ and assume

$$b \Vdash (\forall x \in s_f)(\exists y \in t_f)\phi_f\,.$$

By the definition of the presheaf interpretation we have

$$(\forall u \in m^f)(\exists v \in n)\psi\,.$$

We can now apply Proposition 2.6 and get a set p of subsets of n, independent of f and z, and that there exists w in p such that

$$\mathrm{coll}(u \in m^f, v \in w, \psi)\,. \tag{6.8}$$

For w in p define

$$w' =_{\mathrm{def}} \{(f, g \circ h, y_h) \mid (f, g, y) \in w\,,\ h \in \mathsf{y}(\mathsf{dom}\, f)\}\,.$$

For each w in p we now define a function l_w with domain $\mathsf{y}\,a$ as follows. For f in $\mathsf{y}\,a$, the function $l_w(f)$ is defined by letting, for g' in $\mathsf{y}\,b$

$$l_w(f)(g') =_{\mathrm{def}} \{y' \mid (f, g', y') \in w'\}$$

and observe that $l_w(f)$ is in $V_b^{(\mathbb{C})}$ by assumption that p is a set of subsets of n and the definition of w'. Now define

$$q =_{\mathrm{def}} \{l_w \mid w \in p\}\,.$$

We have just shown that q is a set of functions such that for all u in p and for all f in $\mathsf{y}\,a$, uf is in $V_b^{(\mathbb{C})}$, where $b =_{\mathrm{def}} \mathsf{dom}\, f$. We now wish to conclude the proof by showing that there exists u in q such that

$$b \Vdash \mathrm{coll}(x \in s_f, y \in uf, \phi)$$

Define $u =_{\mathrm{def}} l_w$, where w satisfies (6.8) and define $l =_{\mathrm{def}} uf$. Let g in $\mathsf{y}\,b$ and x in $s(f \circ g)$. We have that there is y such that (f, g, y) is in w and

$$\mathsf{dom}\, g \Vdash \phi_{f \circ g}[z_g/z]$$

by (6.8) and the definition of ψ. Therefore y is in $l(g)$ by definition of l_w and the fact that w is included in w'. Observe that y is in $u(f)(g)$, by the definition of l, as wanted. Now let g' in $\mathsf{y}\,b$ and y' in $l(g')$. Then there are g, h and y such that (f, g, y) is in w and

$$(g \circ h, y_h) = (g', y')\,. \tag{6.9}$$

Using (6.8) and the definition of ψ we obtain that there is x in $s(f \circ g)$ such that

$$\mathsf{dom}\, g \Vdash \phi_{f \circ g}[z_g/z]\,,$$

and therefore

$$\mathsf{dom}\, h \Vdash \phi_{f \circ g \circ h}[x_h, y_h, z_{g \circ h}/x, y, z]$$

by Lemma 6.8. Once we define $x' =_{\mathrm{def}} x_h$, we get

$$\mathsf{dom}\, g' \Vdash \phi_{f \circ g'}[x', y', z_{g'}/x, y, z]$$

and that x' is in $s(f \circ g')$ by (6.9), as wanted. □

Proposition 6.16. *Assuming Subset Collection axiom, the presheaf interpretation of Subset Collection is valid.*

Proof. Let a in $\mathbb{C}_0$, let s and t be in $V_a^{(\mathbb{C})}$, and let ϕ be a formula of $\mathcal{L}^{(a)}$ with $\mathrm{FV}\phi = \{x, y, z\}$. Let q be a set as in the conclusion of Lemma 6.15 and define a function r with domain $\mathsf{y}\, a$ by letting, for f in $\mathsf{y}\, a$

$$rf =_{\mathrm{def}} \{uf \mid u \in q\}$$

Let f in $\mathsf{y}\, a$, define $b =_{\mathrm{def}} \mathsf{dom}\, f$. Let z in $V_b^{(\mathbb{C})}$ and assume

$$b \Vdash (\forall x \in s_f)(\exists y \in b_f)\phi_f$$

Then there exists u in q such that

$$b \Vdash \mathrm{coll}(x \in s_f, y \in uf, \phi[z_f/z])\,.$$

Define $v =_{\mathrm{def}} uf$ and observe that v is in rf, and obviously that

$$b \Vdash \mathrm{coll}(x \in s_f, y \in v, \phi)\,.$$

It is now sufficient to 'close' r under restrictions to conclude the proof, but we leave these details to the reader. □

The next theorem summarises the results obtained so far. Its proof can be obtained working informally in $\mathbf{CZF}^-$.

Theorem 6.17. *Let $\mathbb{C}$ be a small category.*

(i) *The presheaf interpretation of* $\mathbf{CZF}^-$ *in* $V^{(\mathbb{C})}$ *is valid.*

(ii) *Assuming Subset Collection, the presheaf interpretation of* $\mathbf{CZF}$ *in* $V^{(\mathbb{C})}$ *is valid.*

Proof. Part (i) is a consequence of Proposition 6.11, Proposition 6.12 and Proposition 6.14. For part (ii) use part (i) and Proposition 6.16. □

6.7 Sheaf interpretations

We have completed the first step towards sheaf interpretations, which consisted in isolating the presheaf $V^{(\mathbb{C})}$, defining the presheaf interpretation, and proving that all the axioms of **CZF** are valid under the interpretation. In this section, we start to consider the second step of sheaf interpretations. This will consist in defining a reinterpretation of logic, using a syntactic translations and the presheaf interpretation.

The syntactic translation we consider is determined by a presheaf map on the presheaf of sieves, that we defined in Subsection 6.2.3. This presheaf maps satisfy properties analogous to the ones of a nucleus on a set-generated frame, as defined in Subsection 4.2.2. Recall from Subsection 6.2.3 that Ω is defined so that we have

$$\Omega_a = \{p \mid p \text{ set-sieve on } a\}$$

for all a in $\mathbb{C}_0$.

Definition 6.18. Let j be a presheaf map from Ω to Ω. Let us write $j(p)$ instead of $j_a(p)$, for all a in $\mathbb{C}_0$ and all p in Ω_a. We say that j is a **Lawvere-Tierney operator** if for all a in $\mathbb{C}_0$ the following properties

- j is monotone, i.e. $(\forall u, v \in \Omega_a) u \subseteq v \rightarrow ju \subseteq jv$,
- j is inflationary, $(\forall u \in \Omega_a) u \subseteq ju$,
- j is idempotent, i.e. $(\forall u \in \Omega_a) j(ju) \subseteq ju$,
- j respects intersections, i.e. $(\forall u, v \in \Omega_a) ju \cap jv \subseteq j(u \cap v)$,

hold. ◊

6.7.1 Defining sheaf interpretations

From now on, we will consider an fixed, arbitrary presheaf map j and assume that it is a Lawvere-Tierney operator. We will adopt the notation introduced in Definition 6.18, and write $j(p)$ instead of $j_a(p)$, for a in $\mathbb{C}_0$ and p in Ω_a.

To discuss the reinterpretation of logic determined by j, it is convenient to write $\mathcal{L}^{(\mathbb{C})}$ for the language that extends $\mathcal{L}$ with constants for elements of $V_a^{(\mathbb{C})}$, for all a in $\mathbb{C}_0$. The reinterpretations of logic determined by j will be obtained in two steps. The first step consists in introducing a language $\mathcal{L}^{(\mathbb{C},J)}$ that includes $\mathcal{L}^{(\mathbb{C})}$ and has also a modal operator J, and in extending the presheaf interpretation to $\mathcal{L}^{(\mathbb{C},J)}$ using the

Lawvere-Tierney topology j. The second step involves defining a syntactic translation from $\mathcal{L}^{(\mathbb{C})}$ to $\mathcal{L}^{(\mathbb{C},J)}$ using the modal operator J. We will refer to this translation as J-translation. Once we performed these two steps, the sheaf interpretation of a sentence will be simply defined as the presheaf interpretation of its J-translation.

Let us begin by adding two symbols to the language $\mathcal{L}^{(\mathbb{C})}$ and defining the presheaf interpretation for them. We define $\mathcal{L}^{(\mathbb{C},J)}$ to be the extension of the language $\mathcal{L}^{(\mathbb{C})}$ that is obtained by adding a binary relation symbol $=_J$ and a unary modal operator J. It is immediate to define substitutions and restrictions for formulas in the extended language, following the conventions we fixed in Section 6.4. For a in $\mathbb{C}_0$, we write $\mathcal{L}^{(a,J)}$ for the sublanguage of $\mathcal{L}^{(\mathbb{C},J)}$ with constants that are only in $V_a^{(\mathbb{C})}$. We now wish to extend the presheaf interpretation to $\mathcal{L}^{(a,J)}$, for all a in $\mathbb{C}_0$. The necessary ingredient to interpret the binary relation symbol $=_J$ is provided by the next lemma.

Lemma 6.19. *We can define a binary relation* $=_j$ *such that we have*

$$a \Vdash s =_j t \quad \leftrightarrow \quad a \Vdash (\forall x \in s) J (\exists y \in t) x =_j y \wedge (\forall y \in t) J (\exists x \in s) x =_j y$$

for all a *in* $\mathbb{C}_0$ *and all* s, t *in* $V_a^{(\mathbb{C})}$.

Proof. The desired relation can be defined by double set recursion, and recalling the definition of the presheaf interpretation from Table 6.1. □

We now define the presheaf interpretation of the symbol $=_J$ using the relation defined in Lemma 6.19. For a in $\mathbb{C}_0$ and s, t in $V_a^{(\mathbb{C})}$ define

$$a \Vdash s =_J t \quad =_{\mathrm{def}} \quad a \Vdash s =_j t \,. \tag{6.10}$$

We may anticipate that Lemma 6.19 will guarantee us the validity of the sheaf interpretation of Extensionality. The next two definitions and the subsequent lemma will be helpful to define the presheaf interpretation of formulas containing the modal operator J. For a in $\mathbb{C}_0$, p in Ω_a, and ϕ sentence of $\mathcal{L}^{(a)}$ define

$$\begin{aligned} a \Vdash p \quad &=_{\mathrm{def}} \quad p = \mathsf{y}\, a \,, \\ a \Vdash p \rightarrow \phi \quad &=_{\mathrm{def}} \quad (\forall f \in \mathsf{y}\, a)\, \mathsf{dom}\, f \Vdash p_f \rightarrow \mathsf{dom}\, f \Vdash \phi_f \,. \end{aligned}$$

The next lemma simplifies these definitions.

Lemma 6.20. *Let* a *in* $\mathbb{C}_0$. *For* p *in* Ω_a, *and* ϕ *sentence of* $\mathcal{L}^{(a)}$ *the following equivalences*

(i) $a \Vdash p$ *if and only if* $id_a \in p$,

(ii) $a \Vdash p \rightarrow \phi$ *if and only if* $(\forall f \in p)\, \mathsf{dom}\, f \Vdash \phi_f$, *hold.*

Proof. Both (i) and (ii) follow recalling that elements of Ω_a are sieves. $\square$

For ϕ a sentence of $\mathcal{L}^{(a,J)}$ define

$$a \Vdash J\phi =_{\text{def}} (\exists u \in \Omega_a)(a \Vdash ju \wedge a \Vdash u \to \phi) \, .$$

The reader might also gain some intuition into this definition by comparing it with the definition of the extension of a nucleus given in Subsection 5.2.1. The next lemma is intended to make this correspondence more evident.

Lemma 6.21. *Let a be in $\mathbb{C}_0$ and let ϕ be a sentence of $\mathcal{L}^{(a,J)}$. We have*

$$id_a \in [\![J\phi]\!] \leftrightarrow (\exists u \in \Omega_a)\big(id_a \in ju \wedge u \subseteq [\![\phi]\!]\big) \, .$$

Proof. The conclusion follows by unfolding definitions and Lemma 6.20. $\square$

We now come to the definition of the J-translation of $\mathcal{L}^{(\mathbb{C})}$ into $\mathcal{L}^{(\mathbb{C},J)}$. For a formula ϕ of $\mathcal{L}^{(\mathbb{C})}$, we indicate its J-translation with $\langle\phi\rangle$. The J-translation is defined in Table 6.2 by structural induction on formulas.

Table 6.2: Definition of the J-translation.

$$\begin{aligned}
\langle s = t \rangle &=_{\text{def}} s =_J t \, , \\
\langle \bot \rangle &=_{\text{def}} J\bot \, , \\
\langle \phi \wedge \psi \rangle &=_{\text{def}} J\langle\phi\rangle \wedge J\langle\psi\rangle \, , \\
\langle \phi \vee \psi \rangle &=_{\text{def}} J\langle\phi\rangle \vee J\langle\psi\rangle \, , \\
\langle \phi \to \psi \rangle &=_{\text{def}} J\langle\phi\rangle \to J\langle\psi\rangle \, , \\
\langle (\forall x \in s)\phi \rangle &=_{\text{def}} (\forall x \in s) J\langle\phi\rangle \, , \\
\langle (\exists x \in s)\, \phi \rangle &=_{\text{def}} (\exists x \in s)\, J\langle\phi\rangle \, , \\
\langle (\forall x)\, \phi \rangle &=_{\text{def}} (\forall x)\, J\langle\phi\rangle \, , \\
\langle (\exists x)\, \phi \rangle &=_{\text{def}} (\exists x)\, J\langle\phi\rangle \, .
\end{aligned}$$

We now have the two ingredients necessary to define sheaf interpretations: the presheaf interpretation and the J-translation. We simply define the sheaf interpretation as the composition of the J-translation and of the presheaf interpretation. For a sentence ϕ of $\mathcal{L}^{(a)}$, we say that

$$a \Vdash J\langle\phi\rangle$$

is the **sheaf interpretation** of ϕ. Now, recall from Proposition 6.9 that, for a in $\mathbb{C}_0$, restricted sentences of $\mathcal{L}^{(a)}$ correspond to set-sieves on a, i.e. elements of Ω_a. The next lemma compares the action of the modality J on a restricted sentence with the action of the Lawvere-Tierney operator on the set-sieve determined by it.

Lemma 6.22. *Let a be in $\mathbb{C}_0$. For a restricted sentence θ of $\mathcal{L}^{(a,J)}$, we have $a \Vdash J\theta$ if and only if $a \Vdash j[\![\theta]\!]$.*

Proof. The conclusion follows by Lemma 6.20 and Lemma 6.21. □

The next proposition spells out that the modal operator J acts on restricted formulas just as the Lawvere-Tierney operator acts on set-sieves.

Proposition 6.23. *Let a in $\mathbb{C}_0$. For θ, η restricted sentences of $\mathcal{L}^{(a,J)}$, the following properties*

(i) *If $a \Vdash \theta \rightarrow \eta$ then $a \Vdash J\theta \rightarrow J\eta$,*

(ii) $a \Vdash \theta \rightarrow J\theta$,

(iii) $a \Vdash J(J\theta) \rightarrow J\theta$,

(iv) $a \Vdash J\theta \wedge J\eta \rightarrow J(\theta \wedge \eta)$,

hold.

Proof. The claims are direct consequences of the properties of a Lawvere-Tierney operator and of Lemma 6.22. □

We now set out to prove that the properties in Proposition 6.23 hold not only for restricted formulas but for arbitrary ones as well. This is completely analogous to what we did in Subsection 5.2.1, where we showed that, given a nucleus on lower sets, we can extend it to an operator on lower classes that inherits its properties. The following variation over Proposition 2.4 will be useful.

Lemma 6.24. *Let a in $\mathbb{C}_0$. Let p be a subset of $\mathsf{y}\,a$, ψ be a formula of the language $\mathcal{L}$ with free variables f, u and let ϕ be a sentence of the language $\mathcal{L}^{(a,J)}$. Assume that*

$$(\forall f \in p)(\exists u \in \Omega_{\mathsf{dom}\, f})\big(\psi \wedge u \subseteq [\![\phi]\!]_f\big)$$

and

$$(\forall f \in p)(\forall u, v \in \Omega_{\mathsf{dom}\, f})\big((u \subseteq v \subseteq [\![\phi]\!]_f \wedge \psi) \rightarrow \psi[v/u]\big)$$

hold. Then there exists q in Ω_a such that

$$q \subseteq [\![\phi]\!] \wedge (\forall f \in p)\psi[q_f/u]$$

holds.

Proof. We will want to apply Proposition 2.3. Let us define a formula to make this simpler. For f and u define

$$\xi =_{\text{def}} u \in \Omega_{\mathsf{dom}\, f} \wedge u \subseteq [\![\phi]\!]_f \wedge \psi\,.$$

We have

$$(\forall f \in p)(\exists u)\xi$$

by our first assumption. Hence we can apply Proposition 2.3 and obtain that there is a function m with domain p such that

$$(\forall f \in p)\big((\exists u)(u \in m(f)) \wedge (\forall u \in m(f))\xi\big)\,. \tag{6.11}$$

We claim that the set q defined as

$$q =_{\text{def}} \{f \circ g \mid f \in p\,,\ (\exists u \in m(f))g \in u\}$$

satisfies the conclusion of the lemma. We first show that q is in Ω_a. Let $f \circ g$ in q and h in $\mathsf{y}\, c$, where $c =_{\text{def}} \mathsf{dom}\, g$. Hence f is in p and there is u in $m(f)$ such that g is in $m(f)$. We have that u is a sieve by (6.11) and hence $g \circ h$ in u. Therefore $f \circ g \circ h$ is in q, by the definition of q. We therefore have that q is a sieve, and hence it is in Ω_a, as wanted. We now want to show that

$$q \subseteq [\![\phi]\!] \wedge (\forall f \in p)\psi[q_f/u]\,.$$

We begin by showing the first conjunct. Let f is in p, u in $m(f)$ and g is in u, so that $f \circ g$ in q, by the definition of q. By (6.11), u is a subset of $[\![\phi]\!]_f$ and therefore g is in $[\![\phi]\!]_f$. Since $[\![\phi]\!]_f$ is a sieve, we have $f \circ g$ in $[\![\phi]\!]$ and this shows that q is a subset of $[\![\phi]\!]$, as wanted. We now move on and prove the second conjunct. Let f in p. Observe that by (6.11) there exists u in $m(f)$ and we can assume that ξ holds. Therefore we have that both u and q_f are in $\Omega_{\mathsf{dom}\, f}$ and it holds that

$$u \subseteq q_f \subseteq [\![\phi]\!]_f \wedge \psi \, ,$$

where the inclusion $u \subseteq q_f$ is an immediate consequence of the definition of q. The desired conclusion follows by the second assumption in the statement of the lemma. $\square$

It is worth pointing out that in the proof we just presented, we made essential use of Proposition 2.3 and hence of the Strong Collection axiom of $\mathbf{CZF}^-$. The next lemma will be the key to prove the 'idempotency' of the modal operator J.

Lemma 6.25. *Let a in $\mathbb{C}_0$. Let p in Ω_a and let ϕ a sentence of $\mathcal{L}^{(a,J)}$. Assume that*

$$a \Vdash p \rightarrow J\phi \, .$$

Then there exists q in Ω_a such that both $a \Vdash p \rightarrow jq$ and $a \Vdash q \rightarrow \phi$ hold.

Proof. For f in p, u in $\Omega_{\mathsf{dom}\, f}$ we define $\psi =_{\mathrm{def}} \mathsf{y}(\mathsf{dom}\, f) \subseteq ju$. By the assumption, we can apply Lemma 6.24 and get q in Ω_a such that

$$q \subseteq [\![\phi]\!] \wedge (\forall f \in p)\ \mathsf{y}(\mathsf{dom}\, f) \subseteq j(q_f) \, .$$

Therefore we obtain

$$(a \Vdash q \rightarrow \phi) \wedge (a \Vdash p \rightarrow jq)$$

by Lemma 6.20 and the definition the sheaf interpretation, as wanted. $\square$

Proposition 6.26. *Let a be in $\mathbb{C}_0$. Let ϕ and ψ be sentences of $\mathcal{L}^{(a,J)}$. Then the following properties*

(i) *If $a \Vdash \phi \rightarrow \psi$ then $a \Vdash J\phi \rightarrow J\psi$,*

(ii) $a \Vdash \phi \rightarrow J\phi$,

(iii) $a \Vdash J(J\phi) \rightarrow J\phi$,

(iv) $a \Vdash J\phi \wedge J\psi \rightarrow J(\phi \wedge \psi)$,

hold.

Proof. For parts (i), (ii) and (iv) direct calculations suffice. For part (iii) let f in $\mathsf{y}\, a$ and assume $b \Vdash J(J\phi_f)$ where $b =_{\mathrm{def}} \mathsf{dom}\, f$. Then there is p in Ω_b such that

$$(b \Vdash jp) \wedge (b \Vdash p \rightarrow J\phi_f) \qquad (6.12)$$

by definition of the sheaf interpretation. By the second conjunct in (6.12) and Lemma 6.25, there is q in Ω_b such that

$$(b \Vdash p \rightarrow jq) \wedge (b \Vdash q \rightarrow \phi_f)\,.$$

By the first conjunct in (6.12) and the fact that j is monotone and idempotent we obtain $b \Vdash jq$. Therefore we proved

$$(b \Vdash jq) \wedge (b \Vdash q \rightarrow \phi_f)\,,$$

and hence that $b \Vdash J\phi_f$ as required. □

We can now define what it means for the sheaf interpretation of a constructive set theory to be valid.

Definition 6.27. For a in $\mathbb{C}_0$ and ϕ a sentence of $\mathcal{L}^{(a)}$, we say that the sheaf interpretation of ϕ is **valid** if it holds that

$$a \Vdash J\langle\phi\rangle\,.$$

We say that the sheaf interpretation of an axiom scheme of a constructive set theory is **valid** if for all a in $\mathbb{C}_0$ the interpretation of all of the instances of the axiom scheme with parameters in $V_a^{(\mathbb{C})}$ are valid. The sheaf interpretation of a constructive set theory in $V^{(\mathbb{C})}$ is said to be **valid** if the interpretation of all its axioms and axiom schemes are valid. ◊

It seems necessary to assume that, for a in $\mathbb{C}_0$, it holds

$$a \Vdash \neg J\bot\,,$$

in order to validate the axiom of intuitionistic logic concerning negation, and thus in what follows we will work under this assumption.

6.8 Validity of the sheaf interpretation

We now investigate the validity of the sheaf interpretation of the axioms of **CZF**. We consider the axioms of **CZF**$^-$ first and Subset Collection at a later stage. This is because it does not seem possible to prove the validity of Subset Collection without further assumptions on the Lawvere-Tierney topology j. A similar observation applied to Heyting-valued interpretations in Section 5.5.

6.8.1 Structural and set existence axioms

The validity of the sheaf interpretation of all the logical axioms follows by a series of routine verifications. In the following we first consider structural and set existence axioms, and then collection axioms. The next lemma takes care of the structural axioms.

Lemma 6.28. *The sheaf interpretation of Extensionality and Set Induction is valid.*

Proof. The validity of Extensionality is a direct consequence of the definition of the sheaf interpretation of equality and Lemma 6.19. This is indeed obtained by first performing the J-translation, as we defined in Table 6.2, and then applying the presheaf interpretation. Once this is done, the conclusion follows by Lemma 6.19. The validity of the sheaf interpretation of Set Induction follows in a straightforward fashion by the validity of its presheaf interpretation, that was obtained in Section 6.5. □

We now move on to consider set existence axioms.

Proposition 6.29. *The sheaf interpretation of the Pairing, Union, Infinity and Restricted Separation is valid.*

Proof. To prove the validity of these axioms, we invite the reader to observe that the J-translation, as defined in Table 6.2, is analogous to the double-negation translation of classical logic into intuitionistic logic, and that the double-negation satisfies all the properties of the modal operator J as described in Proposition 6.26.

Let us now observe that, for a in $\mathbb{C}_0$ and ϕ a sentence in the language $\mathcal{L}_a^{(\mathbb{C})}$, in order to prove that the sheaf interpretation of ϕ is valid it is sufficient to show that the presheaf interpretation of ϕ is valid and to prove that

$$a \Vdash \phi \rightarrow J\langle\phi\rangle \,. \tag{6.13}$$

We can exploit this fact when considering the sheaf interpretation of the set existence axioms, since we know from Proposition 6.12 that their presheaf interpretation is valid. We therefore need to show only that (6.13) holds where ϕ is an instance of a set existence axioms with parameters in $V_a^{(\mathbb{C})}$. This can be proved very easily, by generalising the proofs concerning the double-negation translation of Pairing, Union, Infinity and Restricted Separation [34, 36]. □

6.8.2 Collection axioms

The next lemma is the key to prove the validity of the sheaf interpretation of Strong Collection. Its statement is inspired by Lemma 5.16, that was used to prove that the Heyting-valued interpretation of Strong Collection is valid.

Lemma 6.30. *Let a in $\mathbb{C}_0$, s in $V_a^{(\mathbb{C})}$, f in $\mathsf{y}\,a$, and let ϕ a formula of $\mathcal{L}^{(a,J)}$ such with* $\mathrm{FV}\phi = \{x, y\}$. *Assume that*

$$b \Vdash (\forall x \in s_f) J (\exists y) \phi_f$$

where $b =_{\mathrm{def}} \mathsf{dom}\, f$. Define

$$m^f =_{\mathrm{def}} \{(g, x) \mid g \in \mathsf{y}(b)\,,\ x \in s(f \circ g)\}$$

and

$$P =_{\mathrm{def}} \{(g, x, y) \mid (g, x) \in m^f\,,\ y \in V^{(\mathbb{C})}_{\mathsf{dom}\, g}\,, \mathsf{dom}\, g \Vdash \phi_{f \circ g}\}$$

Then there exists a subset u of P such that

$$(\forall (g, x) \in m^f)\, \mathsf{y}(\mathsf{dom}\, g) \subseteq j\{h \in \mathsf{y}\,\mathsf{dom}\, g \mid (\exists y)(g \circ h, x_h, y) \in u\}$$

Proof. We will want to apply Proposition 2.4, so for (g, x) in m^f define

$$A_{(g,x)} =_{\mathrm{def}} [\![(\exists y) \phi_{f \circ g}]\!]$$

and for u define

$$\psi =_{\mathrm{def}} u \in \Omega_{\mathsf{dom}\, g} \wedge \mathsf{y}(\mathsf{dom}\, g) \subseteq ju$$

From the assumption of the lemma, using the definition of the sheaf interpretation, we get

$$(\forall (g, x) \in p)(\exists u)(u \subseteq A_{(g,x)} \wedge \psi)$$

Furthermore, since j is monotone and inflationary, we have

$$(\forall (g, x) \in p)(\forall u, v)(u \subseteq v \subseteq A_{(g,x)} \wedge \psi \to \psi[v/u])$$

and therefore we can apply Proposition 2.4. We thus get a function l with domain m^f such that

$$(\forall (g, x) \in m^f) l(g, x) \subseteq A_{(g,x)} \wedge \psi[l(g, x)/u] \tag{6.14}$$

Now define

$$q =_{\mathrm{def}} \{(g \circ h, x_h) \mid g \in \mathsf{y}\, b\,,\ x \in s(f \circ h)\,,\ h \in l(g, x)\}$$

and observe that

$$(\forall (g \circ h, x_h) \in q)(\exists y)\xi$$

where

$$\xi =_{\text{def}} x_h \in s(f \circ g \circ h) \wedge y \in V^{(\mathbb{C})}_{\mathsf{dom}\, h} \wedge \mathsf{dom}\, h \Vdash \phi_{f \circ g \circ h}[x_h/x]$$

We can therefore apply Proposition 2.3 and obtain that there is a function n such that

$$(\forall (g \circ h, x_h) \in q)\big((\exists y \in n(g \circ h, x_h))\xi \wedge (\forall y \in n(g \circ h, x_h))\xi\big) \tag{6.15}$$

To complete the proof, define

$$u =_{\text{def}} \{(g \circ h, x_h, y) \mid (g \circ h, x_h) \in q\,,\ y \in n(g \circ h, x_h)\}$$

The desired conclusion now follows simply unfolding definitions and using (6.14) and (6.15). By means of illustration, we prove that

$$(\forall (g, x) \in m^f)\, \mathsf{y}(\mathsf{dom}\, g) \subseteq j\{h \in \mathsf{y}(\mathsf{dom}\, g) \mid (\exists y)(g \circ h, x_h, y) \in u\}$$

and leave the verification that u is a subset of P to the reader. Let (g, x) in m^f and define $c =_{\text{def}} \mathsf{dom}\, g$. By (6.14) and the definition of ψ we get

$$\mathsf{y}\, c \subseteq j(l(g, x))$$

By the definition of q and the fact that j is monotone we have

$$j(l(g, x)) \subseteq j\{h \mid (g \circ h, x_h) \in q\}$$

By (6.15) and the definition of u we also have

$$\{h \in \mathsf{y}\, c \mid (g \circ h, x_h) \in q\} \subseteq \{h \in \mathsf{y}\, c \mid (\exists y)(g \circ h, x_h, y) \in u\}$$

Combining these inclusions and using that j is monotone and idempotent, we get the desired result. □

Proposition 6.31. *The sheaf interpretation of Strong Collection is valid.*

Proof. Let a, f and ϕ as in the hypothesis of Lemma 6.30 and assume

$$b \Vdash (\forall x \in s_f) J(\exists y)\phi_f$$

where $b =_{\text{def}} \mathsf{dom}\, f$. We then get a set u that satisfies the conclusion of Lemma 6.30. To show the validity of Strong Collection it suffices to find an element t of $V_b^{(\mathbb{C})}$ such that

$$b \Vdash (\forall x \in s_f) J(\exists y \in t)\phi_f \,, \tag{6.16}$$

and

$$b \Vdash (\forall x \in s_f)(\exists y \in t)\phi_f \,. \tag{6.17}$$

We define t as a function with domain $\mathsf{y}\, b$ by letting, for g in $\mathsf{y}\, b$

$$t(g) =_{\text{def}} \{ y \mid (\exists x)(g, x, y) \in u \} \,.$$

Note that t need not be closed under restrictions, and hence need not be in $V_b^{(\mathbb{C})}$, but we can always replace t with another function t' that is so, as we did in the proof of Proposition 6.14. In the following, however, we prefer to assume that t itself is closed under restrictions and leave to the reader the task of adapting the proof where appropriate. We begin by proving (6.16). Let g in $\mathsf{y}\, b$, x in $s(f \circ g)$ and show

$$c \Vdash J(\exists y \in t_g)\phi_{f \circ g} \tag{6.18}$$

where $c =_{\text{def}} \mathsf{dom}\, g$. In order to do so, define

$$q =_{\text{def}} \{ h \in \mathsf{y}\, c \mid (\exists y)(g \circ h, x_h, y) \in u \}$$

and observe that we have $c \Vdash jq$ by the conclusion of Lemma 6.30. Now let h in q and show

$$d \Vdash (\exists y \in t_{g \circ h})\phi_{f \circ g \circ h}$$

where $d =_{\text{def}} \mathsf{dom}\, h$. By definition of q, there is y such that $(g \circ h, x_h, y)$ is in u. By the conclusion of Lemma 6.30, we know that u is a subset of P, as defined in the statement of the lemma, and hence $x_h \in s(f \circ g \circ h)$ and

$$d \Vdash \phi_{f \circ g \circ h}[x_h / x] \,.$$

Furthermore, we have that y is in $t(g \circ h)$, i.e. $t_f(g)$, as wanted. We have therefore found q such that $c \Vdash jq$ and $c \Vdash q \to (\exists y \in t_g)\phi_{f \circ g}$ and therefore we proved (6.18), as wanted. The definition of t and the assumption that u satisfies the conclusion of Lemma 6.30 lead to prove (6.17) without difficulty. □

We now consider Subset Collection. Analogously to what happens when considering Heyting-valued interpretations, the validity of Subset Collection seems to require an extra assumption. When considering Heyting-valued interpretations in Section 5.5, we used the assumption that the frame was set-presented. It seems therefore appropriate to introduce the following notion.

Definition 6.32. We say that a Lawvere-Tierney operator j is **set-presented** there exists a set r such that for all a in $\mathbb{C}_0$ and all p in Ω_a the property

$$f \in jp \leftrightarrow (\exists u \in r)(f \in ju \wedge u \subseteq p)$$

holds. ◊

Let us now assume that j is a set-presented Lawvere-Tierney operator. The next lemma is proved assuming Subset Collection, and it is analogous to Lemma 5.18.

Lemma 6.33. *Let a in $\mathbb{C}_0$, s, t in $V_a^{(\mathbb{C})}$ and let ϕ be a formula of $\mathcal{L}^{(a,J)}$ with* $\mathrm{FV}\phi = \{x, y, z\}$*. Then there exists a set q of functions with domain $\mathsf{y}\, a$ such that*

$$(\forall u \in p)(\forall f \in \mathsf{y}\, a) u(f) \in V^{(\mathbb{C})}_{\mathsf{dom}\, f}$$

and for all f in $\mathsf{y}\, a$ and z in $V^{(\mathbb{C})}_{\mathsf{dom}\, b}$ if

$$b \Vdash (\forall x \in s_f) J(\exists y \in t_f)\phi$$

then there exists u in q such that

$$b \Vdash \langle \mathrm{coll}(x \in s_f, y \in u(f), \phi) \rangle$$

Proof. Since the proof follows the pattern of the proofs of Lemma 5.18 and Lemma 6.15, we leave it to the reader. We limit ourselves to point out that it can be obtained with two applications of Proposition 2.6. □

Proposition 6.34. *The sheaf interpretation of Subset Collection is valid.*

Proof. Let a in $\mathbb{C}_0$, s, t in $V_a^{(\mathbb{C})}$ and let ϕ be a formula of $\mathcal{L}^{(\mathbb{C},J)}$ with x, y, z as free variables. Using a set q that satisfies the conclusion of Lemma 6.33, define a function v with domain $\mathsf{y}\, a$ by letting, for f in $\mathsf{y}\, a$

$$v(f) =_{\mathrm{def}} \{u(f) \mid u \in q\}$$

and observe that for all f in $\mathsf{y}\, a$, we have $v(f) \subseteq V^{(\mathbb{C})}_{\mathsf{dom}\, f}$, by Lemma 6.33. If we let f in

$\mathsf{y}\, a$, define $b =_{\text{def}} \mathsf{dom}\, f$, let z in $V_b^{(\mathbb{C})}$ and assume

$$b \Vdash (\forall x \in s_f) J (\exists y \in t_f) \phi \,,$$

then, by Lemma 6.33, we get u in q such that

$$b \Vdash \langle \text{coll}(x \in s_f, y \in u(f), \phi) \rangle \,.$$

Hence we have found $u(f) \in s_f(id_b)$ that satisfies the desired conclusion. Now, observe that v need not be closed under restrictions, but we can always replace it with a function v' that is closed under restriction and such that for all f in $\mathsf{y}\, a$, $v(f) \subseteq v'(f)$, so that the desired claim holds. ☐

We have therefore finally reached the following result.

Theorem 6.35. *Let j be a Lawvere-Tierney operator.*

(i) *The sheaf interpretation of* $\mathbf{CZF}^-$ *is valid.*

(ii) *Assuming the Subset Collection axiom scheme, if j is set-presented, the sheaf interpretation of* **CZF** *is valid.*

Proof. Part (i) follows by Lemma 6.28, Proposition 6.29 and Proposition 6.31. For part (ii), use part (i) and Proposition 6.34. ☐

6.9 Concluding remarks

The interpretations described in this chapter offer a plethora of potential applications, that we regrettably had not the opportunity to develop yet. Two particular kinds of presheaf interpretations arise by considering special examples of small categories: posets and monoids. The first kind leads us to extend the well-known Kripke interpretation for intuitionistic logic to CST. The second kind, instead, does not seem to have been considered yet and represents an interesting direction for further research. Remarkable applications arise in the case in which the category is neither a poset not a monoid. For example, presheaf interpretations have been used to prove the independence of the so-called ‘world’s weakest axiom of choice’ [13, 33].

So far, we have only discussed the first step of the sheaf interpretations. Considering Lawvere-Tierney operators adds further generality and leads to more applications. Lawvere-Tierney operators can be defined starting from a **site**, i.e. a small category equipped with a coverage [49, Section C.2.1]. The notion of a site generalises the one of a posite that we introduced in Section 4.5. Inductive definitions can indeed be used to generate set-presented Lawvere-Tierney topologies from a site, just as we defined a set-presented nucleus from a posite. Remarkably, the role of inductive definitions becomes evident once some of the requirements that are usually part of the notion of a 'Grothendieck topology', such as the 'transitivity axiom' [56, Section III.2] are dropped, as explained in [49, Section C.2.1] As an example of the applications that are allowed by considering Lawvere-Tierney topologies, Dana Scott indicated how coverages on posets can be used to extend Beth interpretations to set theories [78].

Finally, we expect that the second step of the sheaf interpretations described in this chapter leads to a variation of the double-negation translation that is suitable for CST. Such a translation should reduce $\mathbf{CZF}^{-} + \mathbf{REM}$ to $\mathbf{CZF}^{-}$. The definition of the translation should be based on the definition of an operator J that 'lifts' the double-negation operator, in a way similar to one adopted to extend the nucleus from subsets to subclasses in Section 5.2 or to define the presheaf interpretation of the modal operator J in Section 6.7. The translation should then coincide with the standard one for restricted formulas, but not for arbitrary ones, in general. This idea seems essential to derive the translation of Strong Collection within a constructive set theory. The usual definition of the double-negation translation seems indeed to force on us an application of Full Separation to obtain the double-negation translation of Collection [34]. It seems remarkable that that this idea does not seem to have been considered before, but stems very naturally from our development of sheaf interpretations for CST.

Part III

Collection Principles in DTT

Chapter 7

Logic-enriched type theories

7.1 Collection principles

From now on, we focus our attention on Dependent Type Theory (DTT). In particular, we investigate the possibility of transferring in DTT the development of sheaf interpretations obtained for CST. Let us begin by observing that the collection axioms of CST played a double role in Part II. As an inspection of the relevant proofs may reveal, collection axioms were essential both to define sheaf interpretations and to prove the validity of their instances under the interpretation. This simple observation represents a stimulating starting point to investigate sheaf interpretations in DTT, as it suggests to isolate in DTT principles that correspond to the collection axioms of CST as a preliminary step to develop sheaf interpretations in CST.

The type-theoretic interpretation of CST does not help in this respect [2, 3, 4]. The validity of the collection axioms follows indeed by the type-theoretic axiom of choice that is justified by the propositions-as-types idea [59, pages 50 – 52]. Yet, it does not seem appropriate to consider the type-theoretic axiom of choice as the counterpart of the collection axioms of CST. As categorical logic suggests, choice principles are not generally preserved by sheaf interpretations [62]. Avoiding the propositions-as-types treatment of logic seems therefore a useful preliminary step to isolate collection principles in DTT.

We are therefore led to introduce **logic-enriched type theories**, that are extensions of pure type theories in which a primitive treatment of logic is possible. As we will see, in logic-enriched type theories we can formulate type-theoretic principles inspired by the collection principles of CST. There are at least two criteria to test whether these principles correspond exactly to the collection axioms of CST. Firstly, these rules should lead to the formulation of a logic-enriched type theory that is mutually interpretable with the constructive set theory **CZF**, in which both the Strong Collection and the Subset Collection axioms are assumed. Secondly, they should not only be justified by the constructive treatment of logic, as informed by the propositions-as-types idea, but also by the other treatments of logic, as inspired by sheaf interpretations.

As we discussed in Section 1.5, the formulation of collection principles in logic-enriched type theories is not a straightforward task, because of the intrinsic differences between the settings of set theory and type theory. For example, in set theory the language is untyped and there is a clear distinction between sets and classes, while in type theory we have a rich type structure and no explicit notion of 'class'. Furthermore, Extensionality has always been one of the axioms for the set theories we consider, while extensional principles are generally not assumed in the more recent formulations of dependent type theories [65].

Logic-enriched type theories provide us with ways of overcoming all these difficulties. For instance, we may think of propositions depending on a type as 'classes'. Universe types then suggest us a notion of 'smallness' that is reminiscent of the one used in AST to distinguish sets from classes [61, 62]. Once these notions have been formally set up, the introduction of an appropriate notation, partially inspired by the original type-theoretic interpretation of CST, allows us to overcome the difficulties associated to the non-extensionality of dependent type theories. Collection principles can then be formulated as type-theoretic rules, and thus we are led to consider a logic-enriched type theory, called **ML(CZF)**, that includes them.

The logic-enriched type theory **ML(CZF)** will be our focus in the remainder of the thesis. We will indeed show that the collection rules formulated in this chapter satisfy both of the criteria discussed earlier. Apart from the formal connection between constructive set theories and logic-enriched type theories, much of the notation and of the results we are going to introduce and prove in this chapter have been suggested by analogous abbreviations and facts in the context of CST. We hope that readers who are not familiar with DTT may use this informal correspondence to navigate this final part of the thesis.

7.2 Adding logic

A **logic-enriched type theory** has forms of judgement $(\Gamma)\ \mathcal{B}$ where Γ is a context and $\mathcal{B}$ is a judgement body that has either one of the forms allowed in a pure type theory or one of the following:

$$\phi : \mathsf{prop}\,,$$
$$\phi_1, \ldots, \phi_m \Rightarrow \phi\,.$$

These judgements express that, in context Γ, ϕ is a well-formed proposition expression and that ϕ is a logical consequence of the propositions $\phi_1, \ldots, \phi_m$, respectively. The well-formedness of a judgement $\phi_1, \ \ldots, \ \phi_m \ \Rightarrow \ \phi$ in context Γ presupposes not only that Γ is well-formed but also that the judgements $\phi_i \ : \ \mathsf{prop}$ are derivable, for $i = 1, \ldots, m$.

A convention. We will write $(\Gamma)\ \phi$ rather than $(\Gamma)\ \ \Rightarrow \phi$. Recalling from Subsection 3.1.1 that we omit the empty context, we will write ϕ instead of the judgement $\Rightarrow \phi$ for brevity.

The instances of the rules of a logic-enriched type theory have the same form of the one of a pure type theory, i.e.

$$\frac{J_1 \quad \cdots \quad J_n}{J\,,}$$

but the judgements can have all the forms allowed in a logic-enriched type theory.

A convention. As always, in the statement of formation rules we suppress a context that is common to the premisses and conclusion. In the inference rules we will also suppress a list of assumptions appearing on the left hand side of the symbol $\Rightarrow$ in the logical premisses and conclusion.

7.2.1 Extending the raw syntax

The raw syntax of logic-enriched type theories extends the one of pure type theories given in Subsection 3.1.2 with an extra category of

- proposition expressions, (i.e. formulas).

We will refer to formulas as **2-expressions**. The raw expressions of a logic-enriched type theory can be formed according to the same rules given in Subsection 3.1.2, but allowing constant symbols of the signature to have arities $(n_1^{\epsilon_1} \cdots n_k^{\epsilon_k})^{\epsilon}$ where $k \geq 0$, $n_1, \ldots, n_k \geq 0$ and each of $\epsilon, \epsilon_1, \ldots, \epsilon_k$ can be not only 0 or 1, but also 2. We will exploit the raw syntax for logic-enriched type theories in Section 8.6 where we define types-as-classes interpretations.

Recall from Subsection 3.1.2 that raw expressions need not to be well-formed. For formulas, it is the judgement $(\Gamma)\ \phi : \mathsf{prop}$ that is used to express that, in context Γ, ϕ is a well-formed formula.

7.3 Rules for logic-enriched type theories

One of the reasons for the interest in logic-enriched type theories is that they allow a great deal of flexibility in the formulation of rules concerning the judgements expressing logic. We now overview the rules for logic-enriched type theories that are considered in the rest of the thesis.

7.3.1 Predicate logic

Predicate logic rules express formation and inference rules for the intuitionistic logical constants. The forms of proposition that we consider are the usual ones: we use $\top$ and $\bot$ for the canonical true and false propositions, respectively; we then adopt $\wedge, \vee$ and $\supset$ for conjunction, disjunction and implication, respectively. Finally, for each type A of the standard pure type theory that is being considered, we write $(\forall x : A)$ and $(\exists x : A)$ for the universal and existential quantifiers over A. The formation, introduction and elimination rules for these forms of proposition are straightforward. By means of illustration, we present the inference rules for disjunction and existential quantification in Table 7.1 and Table 7.2.

Table 7.1: Rules for disjunction.

$\vee$-formation

$$\frac{\phi_1 : \mathsf{prop} \qquad \phi_2 : \mathsf{prop}}{\phi_1 \vee \phi_2 : \mathsf{prop}}$$

$\vee$-introduction

$$\frac{\phi_1 : \mathsf{prop} \qquad \phi_2 : \mathsf{prop} \qquad \phi_1}{\phi_1 \vee \phi_2} \qquad\qquad \frac{\phi_1 : \mathsf{prop} \qquad \phi_2 : \mathsf{prop} \qquad \phi_2}{\phi_1 \vee \phi_2}$$

$\vee$-elimination

$$\frac{\phi_1 \vee \phi_2 \qquad \phi_1 \Rightarrow \psi \qquad \phi_2 \Rightarrow \psi}{\psi}$$

If $\mathbf{T}$ is a pure type theory then we write $\mathbf{T} + \mathbf{IL}$ for the logic-enriched type theory that is obtained from $\mathbf{T}$ by adding predicate logic rules. We have usual definitions of logical equivalence and negations: for ϕ, ψ define

$$\begin{aligned} \phi \equiv \psi \quad &=_{\text{def}} \quad (\phi \supset \psi) \wedge (\psi \supset \phi)\,, \\ \neg\phi \quad &=_{\text{def}} \quad \phi \supset \bot\,. \end{aligned}$$

7.3.2 Induction rules

It is natural to extend a logic-enriched type theory with additional non-logical rules to express properties of the various forms of type. For instance, we may consider adding a rule for mathematical induction to the rules concerning the type of natural

Table 7.2: Rules for existential quantification.

∃-formation

$$\frac{A : \mathsf{type} \qquad (x : A)\ \phi : \mathsf{prop}}{(\exists x : A)\ \phi : \mathsf{prop}}$$

∃-introduction

$$\frac{a : A \quad (x : A)\ \phi : \mathsf{prop} \qquad \phi[a/x]}{(\exists x : A)\phi}$$

∃-elimination

$$\frac{(\exists x : A)\phi \qquad \psi : \mathsf{prop} \qquad (x : A)\ \phi \Rightarrow \psi}{\psi}$$

numbers and there are similar rules for the other inductive forms of type. Remarkably, induction rules can be described in a uniform way. For each inductive type $C : \mathsf{type}$ of a logic-enriched type theory $\mathbf{T} + \mathbf{IL}$ we have the following induction rule

$$\frac{(z : C)\ \phi : \mathsf{prop} \qquad e : C \qquad \text{Premisses}}{\phi[e/z]}$$

where the premisses of the rule depend on the form of the inductive type C, in the way described in Table 7.3, where we consider all the inductive types of the logic-enriched type theory $\mathbf{MLW} + \mathbf{IL}$.

Table 7.3: Inductive types of **MLW** and premisses of their induction rule.

C	Premisses
$\mathsf{0}$	
$\mathsf{1}$	$\phi[0_1/z]$,
$\mathsf{2}$	$\phi[1_2/z]$, $\quad \phi[2_2/z]$,
N	$\phi[0/z]$, $\quad (x : \mathsf{N})\ \phi[x/z] \Rightarrow \phi[\mathsf{succ}(x)/z]$,
$(\Sigma x : A)B$	$(x : A, y : B)\ \phi[\mathsf{pair}(x, y)/z]$,
$(Wx : A)B$	$(x : A, u : B \to C)\ (\forall y : B)\phi[\mathsf{app}(u, y)/z] \Rightarrow \phi[\mathsf{sup}(x, u)/z]$.

If $\mathbf{T} + \mathbf{IL}$ is a standard logic-enriched type theory, we write $\mathbf{T} + \mathbf{IL} + \mathbf{IND}$ for the standard logic-enriched type theory that is obtained from $\mathbf{T} + \mathbf{IL}$ by adding the induction rules for all the inductive types of $\mathbf{T} + \mathbf{IL}$.

7.4 A proposition universe

Recall from Subsection 3.3.2 that the pure type theory $\mathbf{ML_1}$ has rules for a type universe U reflecting all the forms of type of the pure type theory **ML**. When adding logic to a pure type theory **T** that includes $\mathbf{ML_1}$ it is natural to match the type universe U by adding a proposition universe P. The formation, elimination and introduction rules for the proposition universe P are given in Subsection B.2.4. The reader may observe that they are analogous to the ones for the type universe U given in Subsection B.2.3. In particular, we have the elimination rule

$$\frac{p : \mathsf{P}}{\tau\, p : \mathsf{prop}\,.}$$

For $\phi : \mathsf{prop}$ we say that ϕ is **small** if, for some $p : \mathsf{P}$ the judgement

$$\phi \equiv \tau\, p$$

is derivable, and in that case we say that p is a **representative** for ϕ. The rules for P express that its elements are representatives for propositions whose quantifiers range over small types.

Observe that the computation rules for the type universe U given in in Subsection B.2.3 adopt the equality form of judgement to express that introduction rules of U reflect type formation rules. For the proposition universe P it seems convenient to avoid the use of an equality form of judgement for propositions in order to express that P reflects logic and rather formulate the computation rules for P with logical equivalence. This is useful, for example, to define the types-as-classes interpretation of logic-enriched type theories, as we will do in Section 8.6.

7.5 Collection rules

To present the collection rules corresponding to the collection axioms of **CZF**, we introduce some notation that makes this correspondence more intuitive. For $A : \mathsf{type}$ define $\mathsf{Sub}\, A$, the **type of small subclasses** of A, as follows

$$\mathsf{Sub}\, A =_{\mathrm{def}} (\Sigma x : \mathsf{U})\big((x \to \mathsf{P}) \times (x \to A)\big) \;\; : \mathsf{type}\,.$$

Some intuition. For $e : \mathsf{U}\,, d : e \to \mathsf{P}$ and $f : e \to A$, we may think of a canonical element

$$\mathsf{pair}(e, \mathsf{pair}(d, f)) : \mathsf{Sub}\, A$$

as the ‘class’ of all $\mathsf{app}(f, x)$ for $x : \mathsf{T}\, e$ such that $\tau(\mathsf{app}(d, x))$.

To support this intuition for arbitrary elements of $\mathsf{Sub}\, A$ and not just the canonical ones, we define some expressions, using the projections for elements of Σ-types defined in Subsection 3.3.1. For $a : \mathsf{Sub}\, A$ define

$$\dot{\mathsf{el}}\, a =_{\mathrm{def}} a.1 : \mathsf{U}\,,$$
$$\mathsf{el}\, a =_{\mathrm{def}} \mathsf{T}(\dot{\mathsf{el}}\, a) : \mathsf{type}\,.$$

For $x : \mathsf{el}(a)$ define

$$\dot{\mathsf{dom}}(a, x) =_{\mathrm{def}} \mathsf{app}(a.2.1, x) : \mathsf{P}\,,$$
$$\mathsf{dom}(a, x) =_{\mathrm{def}} \tau(\dot{\mathsf{dom}}(a, x)) : \mathsf{prop}\,,$$

and finally

$$\mathsf{val}(a, x) =_{\mathrm{def}} \mathsf{app}(a.2.2, x) : A\,.$$

Some intuition. An element $a : \mathsf{Sub}\, A$ may be thought of as the ‘class’ of all $\mathsf{val}(a, x)$ for $x : \mathsf{el}\, a$ such that $\mathsf{dom}(a, x)$.

The notation we just introduced allows us to define ‘restricted quantification’ over an element of $\mathsf{Sub}\, A$, as we will see in the next definitions. These definitions will play an essential role in the formulation of the collection rules, and throughout the remainder of the thesis. In order to maintain the distinction between small and arbitrary propositions, the definition of ‘restricted quantification’ has two clauses, depending on whether we have a small or an arbitrary proposition. For each clause, we obviously have both universal and existential quantification.

We begin with small propositions. For $a : \mathsf{Sub}\, A$ and $(x : A)\; p : \mathsf{P}$ define

$$(\dot{\forall} x \in a) p =_{\mathrm{def}} (\dot{\forall} x : \dot{\mathsf{el}}\, a)\, \dot{\mathsf{dom}}(a, x) \mathbin{\dot{\supset}} p[\mathsf{val}(a, x)/x] : \mathsf{P}\,, \tag{7.1}$$
$$(\dot{\exists} x \in a) p =_{\mathrm{def}} (\dot{\exists} x : \dot{\mathsf{el}}\, a)\, \dot{\mathsf{dom}}(a, x) \mathbin{\dot{\wedge}} p[\mathsf{val}(a, x)/x] : \mathsf{P}\,.$$

We now extend this to arbitrary propositions. For $(x : A)\; \phi : \mathsf{prop}$ define

$$(\forall x \in a) \phi =_{\mathrm{def}} (\forall x : \mathsf{el}\, a)\, \mathsf{dom}(a, x) \supset \phi[\mathsf{val}(a, x)/x] : \mathsf{prop}\,, \tag{7.2}$$
$$(\exists x \in a) \phi =_{\mathrm{def}} (\exists x : \mathsf{el}\, a)\, \mathsf{dom}(a, x) \wedge \phi[\mathsf{val}(a, x)/x] : \mathsf{prop}\,.$$

The definition that we introduce next will be very useful to present collection rules. For $A, B : \mathsf{type}$, $a : \mathsf{Sub}\, A$, $b : \mathsf{Sub}\, B$ and $(x : A, y : B)\ \phi : \mathsf{prop}$ define

$$\mathsf{coll}(a, b, (x, y)\phi) =_{\text{def}} (\forall x \in a)(\exists y \in b)\phi \wedge (\forall y \in b)(\exists x \in a)\phi : \mathsf{prop}\,.$$

Observe the correspondence between this abbreviation and the one used in Section A.4 to formulate the collection axioms in set theory. The collection rules are presented in Table 7.4 on page 150. They will allow us to formulate a logic-enriched type theory **ML**(**CZF**) that will be the focus of our investigations for much of Chapter 8 and Chapter 9. The definition of the type theory **ML**(**CZF**) is given in the next section.

Table 7.4: Collection rules for logic-enriched type theories.

$$\frac{A, B : \mathsf{type} \qquad a : \mathsf{Sub}\, A \qquad (x : A, y : B)\ \phi : \mathsf{prop}}{(\forall x \in a)(\exists y : B)\phi \Rightarrow (\exists v : \mathsf{Sub}\, B)\mathsf{coll}(a, v, (x, y)\phi)}$$

$$\frac{A, B, C : \mathsf{type} \qquad a : \mathsf{Sub}\, A \qquad b : \mathsf{Sub}\, B \qquad (x : A, y : B, z : C)\ \psi : \mathsf{prop}}{(\exists u : \mathsf{Sub}(\mathsf{Sub}\, B))(\forall z : C)\big((\forall x \in a)\ (\exists y \in b)\psi \supset (\exists v \in u)\mathsf{coll}(a, v, (x, y)\psi)\big)}$$

7.6 Some logic-enriched type theories

We now formulate logic-enriched type theories that will be considered in the rest of the thesis, and in particular the logic-enriched **ML**(**CZF**) that will be shown to be mutually interpretable with the constructive set theory **CZF** in Chapter 8.

7.6.1 $\mathbf{ML_1} + \mathbf{IL_1}$

The logic-enriched type theory $\mathbf{ML} + \mathbf{IL}$ is obtained from the pure type theory **ML** presented in Subsection 3.3.1 by adding predicate logic rules. The logic-enriched type theory $\mathbf{ML_1} + \mathbf{IL_1}$ extends the logic-enriched type theory $\mathbf{ML} + \mathbf{IL}$ with rules for a type universe and a proposition universe.

7.6.2 ML(CZF)

Recall from Subsection 3.3.4 that $\mathbf{ML_1^-} + \mathbf{W^-}$ is a pure type theory with only restricted version of Π-types and W-types, so that the types of $\mathbf{ML_1^-} + \mathbf{W^-}$ have the following forms

$$\mathsf{0}\,,\ \mathsf{1}\,,\ \mathsf{2}\,,\ \mathsf{N},\ (\Sigma x : A)B\,,\ (\Pi^{-} x : A)B\,,\ (W^{-} x : A)b\,,\ \mathsf{U}\,.$$

To formulate **ML**(**CZF**) we first of all extend this pure type theory with extra W^{-}-elimination rules that allow definitions by 'double recursion' over W^{-}-types, as given in Section B.3. These rules will be used to define a type-theoretic counterpart of the extensional equality relation of **CZF**, and do not seem derivable without the assumption of Π-rules.

The logic-enriched type theory **ML**(**CZF**) is then obtained by adding predicate rules, P-rules, induction rules for all the inductive types of $\mathbf{ML_1^-} + \mathbf{W}^-$, and both the Strong Collection and the Subset Collection rule. The induction rules for W^{-}-types have the same form of the ones for W-types, as given in Table 7.3. The logic-enriched type theory **ML**($\mathbf{CZF}^-$) is obtained from **ML**(**CZF**) by omitting the Subset Collection rule.

Chapter 8

The generalised type-theoretic interpretation of CST

8.1 Iterative small classes

The main aim of this chapter is to define an interpretation of **CZF** into **ML**(**CZF**). There are at least two motivations for this: firstly we want to establish a formal connection between collection axioms and the collection rules formulated in Section 7.5, and secondly we wish to explore whether we can generalise results from pure to logic-enriched type theories, avoiding the propositions-as-types treatment of logic. We are also going to recall how this generalised generalised type-theoretic interpretation can be combined with an interpretation in the reverse direction to show that **CZF** and **ML**(**CZF**) are mutually interpretable [6].

The original type-theoretic interpretation of **CZF** was obtained introducing a pure type theory $\mathbf{ML_1V}$ with rules for a special W-type V, but no rules for arbitrary W-types [2, 3, 4]. This interpretation rests on two main components: the rules for the type V, that is used to interpret the sets of **CZF**, and the propositions-as-types idea, that is adopted to interpret the language of CST. For example, the propositions-as-types idea plays a crucial role to prove the validity of Strong Collection and Subset Collection, as they are proved using the type-theoretic axiom of choice. It is important to observe that these two components are closely intertwined, since the rules for the type V provide a suitable interpretation for the sets of **CZF** only under the assumption that propositions are interpreted as types. This is particularly evident, for example, if one inspects the proof of the validity of Restricted Separation [3, Theorem 2.8].

To define a generalised type-theoretic interpretation of **CZF** we therefore need to reconsider the definition of the type that is used to interpret sets. The definition we are going to adopt becomes more significant recalling that, in set theory, the cumulative hierarchy is defined inductively as the smallest class X such that

$$(\forall x \in \mathsf{Pow}\, X)\; x \in X\,. \tag{8.1}$$

Following the intuition suggested in Section 7.5, we may view types of small subclasses as the counterpart in type theory of power classes in set theory. For $X : \mathsf{type}$, we may indeed consider $a : \mathsf{Sub}\, X$ as a 'subset' of X. This suggests to introduce a type V as the 'smallest' type X for which we can define

$$(x : \mathsf{Sub}\, X)\;\; \mathsf{set}\, x : X\,,$$

that expresses (8.1) in type theory. Such a type V can indeed be defined explicitly in **ML**(**CZF**) using the restricted forms of W-types that are available in it, as we will do in Section 8.3. Using the abbreviations introduced in Section 7.5, we may already think of $\mathsf{set}\, a : \mathsf{V}$, for $a : \mathsf{Sub}\, \mathsf{V}$, as the 'set' all $\mathsf{val}(a, x) : \mathsf{V}$, for $x : \mathsf{el}\, a$ such that $\mathsf{dom}(a, x)$. We will refer to V as the type of **iterative small classes**.

The generalised type-theoretic interpretation will then be defined in analogy with the original one, but here formulas and restricted formulas of **CZF** correspond to propositions and small propositions rather than to types and small types, as in the original interpretation. The validity of Restricted Separation will follow without assuming the propositions-as-types treatment of logic, and the validity of collection axioms will be a consequence of the collection rules and not of the type-theoretic axiom of choice.

We have discussed the issue of interpreting **CZF** into type theories. It is also natural to consider translations in the opposite direction. For example we may interpret types as sets and objects of a type as elements of the corresponding set. Then to each type forming operation there is the natural set forming operation that corresponds to it. In this way we get a conceptually very simple set theoretical interpretation of the type theory **ML**, which has no universes or W-types, in **CZF**, and this extends to an interpretation of **MLW** in $\mathbf{CZF^+}$ and of $\mathbf{MLW_1}$ in an extension $\mathbf{CZF^+u}$ of $\mathbf{CZF^+}$ expressing the existence of a universe in the sense of [67]. The syntactic details of this kind of translation can be found in [5].

A weakness of these types-as-sets interpretations, when linked with the original type-theoretic interpretation, is that there seems to be a mismatch between the set theories and the type theories. So although we get a translation of **CZF** into $\mathbf{ML_1V}$ we only seem to get a translation of **ML** into **CZF** and to translate $\mathbf{ML_1V}$ into a constructive set theory using types-as-sets we seem to need to go to the set theory $\mathbf{CZF^+u}$, which is much stronger than **CZF**. This mismatch is overcome in [5] by having axioms for an infinite hierarchy of universes on both the type theory side and the set theory side. This allows for the two sides to catch up with each other.

A particularly good aspect of the generalised type-theoretic interpretation is that it leads to formulate a straightforward types-as-classes interpretation in the reverse direction and thus show that **CZF** and **ML**(**CZF**) are mutually interpretable. The main reason for considering restricted versions of Π-types and W-types is indeed to make the types-as-classes interpretation possible. Remarkably, these restrictions still allow us to define the generalised type-theoretic interpretation.

8.2 Small subclasses

From now on we work informally within **ML**(**CZF**). In this section, however, we never make use of collection rules. The considerations in Section 8.1 suggest to start our development of the generalised type theoretic interpretation of **CZF** by studying the properties of the type of small subclasses of a type. For some readers, the intuitive analogy between the type of small subclasses of a type and the class of subsets of a class may provide some insight into the series of lemmas 8.1 – 8.6. In these lemmas, we make frequent and sometimes implicit use of the convenient notation for 'restricted quantifiers' introduced in (7.1) and (7.2) on page 149.

Proposition 8.1 (Emptyset). *For* $A : \mathsf{type}$ *we can define*

$$\emptyset : \mathsf{Sub}\, A$$

such that if $(x : A)\ \phi : \mathsf{prop}$ *and* $c = \emptyset : \mathsf{Sub}\, A$ *then the judgements*

$$\begin{aligned} (\forall x \in c)\phi &\equiv \top\,, \\ (\exists x \in c)\phi &\equiv \bot\,, \end{aligned}$$

are derivable.

Proof. Define

$$\begin{aligned} e &=_{\text{def}} \dot{\mathsf{0}} && : \mathsf{U}\,, \\ d &=_{\text{def}} (\lambda\_ : e)\dot{\bot} && : e \to \mathsf{P}\,, \\ f &=_{\text{def}} (\lambda z : e)\, \mathsf{r}_0(z) && : e \to A\,, \end{aligned}$$

and finally

$$\emptyset =_{\text{def}} \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\, A\,.$$

The desired conclusion is an immediate consequence of the definition of d. □

Proposition 8.2 (Pairing). *For* $A : \mathsf{type}$ *we can define*

$$(x, y : A)\ \{x, y\} : \mathsf{Sub}\, A$$

such that if $(x : A)\ \phi : \mathsf{prop}$, $a, b : A$ *and* $c = \{a, b\} : \mathsf{Sub}\, A$ *then the judgements*

$$\begin{aligned} (\forall x \in c)\phi &\equiv \phi[a/x] \wedge \phi[b/x]\,, \\ (\exists x \in c)\phi &\equiv \phi[a/x] \vee \phi[b/x]\,, \end{aligned}$$

are derivable.

Proof. Define

$$\begin{aligned} & e &=_{\mathrm{def}}&\ \dot{\mathsf{2}} && : \mathsf{U}\,, \\ & d &=_{\mathrm{def}}&\ (\lambda\_ : e)\dot{\top} && : e \to \mathsf{P}\,, \\ (x, y : A)\ & f &=_{\mathrm{def}}&\ (\lambda z : e)\, \mathsf{r_2}(x, y, z) && : e \to A\,, \end{aligned}$$

and finally

$$(x, y : A) \quad \{x, y\} \quad =_{\mathrm{def}} \quad \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\, A\,.$$

Let $(x : A)\ \phi : \mathsf{prop}$, $a, b : A$ and assume that $c = \{a, b\} : \mathsf{Sub}\, A$. By the definition of $\{a, b\}$ we have

$$\begin{aligned} & \mathsf{el}\{a, b\} &=&\ \mathsf{2} && : \mathsf{type}\,, \\ (x : \mathsf{2})\ & \mathsf{dom}(\{a, b\}, x) &\equiv&\ \top\,, \\ (x : \mathsf{2})\ & \mathsf{val}(\{a, b\}, x) &=&\ \mathsf{r_2}(a, b, x) && : A\,. \end{aligned}$$

We can obtain the desired conclusion

$$(\forall x \in c)\phi \equiv \phi[a/x] \wedge \phi[b/x]$$

in two steps. First unfold the definitions of restricted quantifiers and get

$$(\forall x \in c)\phi \equiv (\forall x : \mathsf{2})\phi[\mathsf{r_2}(a, b, x)/x]\,.$$

Then, observe that the desired equivalence follows once we prove that

$$(\forall x : \mathsf{2})\phi[\mathsf{r_2}(a, b, x)/x] \equiv \phi[a/x] \wedge \phi[b/x]\,. \tag{8.2}$$

The 'left-to-right' implication in (8.2) is obtained with the ∀-elimination, and the 'right-to-left' implication follows by the 2-induction rule. Following an analogous reasoning we can also prove

$$(\exists x \in c)\phi \equiv \phi[a/x] \vee \phi[b/x]\,.$$

We leave the details to the reader, and limit ourselves to suggest to first unfold the definitions of restricted quantifiers, and then to use the 2-induction rule and the ∨-elimination rule. □

Corollary 8.3 (Singleton). *For* $A : \mathsf{type}$ *we can define*

$$(x, y : A)\ \{x\} : \mathsf{Sub}\, A$$

such that if $(x : A)\ \phi : \mathsf{prop}$*,* $a : A$ *and* $c = \{a\} : \mathsf{Sub}\, A$ *then the judgements*

$$(\forall z \in c)\phi \quad \equiv \quad \phi[a/x]\,,$$

$$(\exists x \in c)\phi \quad \equiv \quad \phi[a/x]\,,$$

are derivable.

Proof. Once we define

$$(x : A)\ \{x\} =_{\mathrm{def}} \{x, x\} : \mathsf{Sub}\, A\,,$$

the conclusion follows by Proposition 8.2. □

Proposition 8.4 (Union). *For* $A : \mathsf{type}$ *we can define*

$$(x : \mathsf{Sub}(\mathsf{Sub}\, A)) \quad \bigcup x \quad : \quad \mathsf{Sub}\, A$$

such that if $(x : A)\ \phi : \mathsf{prop}$*,* $a : \mathsf{Sub}(\mathsf{Sub}\, A)$ *and* $c = \bigcup a : \mathsf{Sub}\, A$ *then the judgements*

$$(\forall x \in c)\phi \quad \equiv \quad (\forall y \in a)(\forall x \in y)\phi\,,$$

$$(\exists x \in c)\phi \quad \equiv \quad (\exists y \in a)(\exists x \in y)\phi\,,$$

are derivable.

Proof. For $x : \mathsf{Sub}(\mathsf{Sub}\, A)$ define

$$\begin{aligned} e &=_{\mathrm{def}} (\dot\Sigma y : \dot{\mathsf{el}}\, x)\, \dot{\mathsf{el}}(\mathsf{val}(x,y)) && : \mathsf{U}\,, \\ d &=_{\mathrm{def}} (\lambda z : e)\, \dot{\mathsf{dom}}(x, z.1) \mathbin{\dot\wedge} \dot{\mathsf{dom}}(\mathsf{val}(x, z.1), z.2) && : e \to \mathsf{P}\,, \\ f &=_{\mathrm{def}} (\lambda z : e)\, \mathsf{val}(\mathsf{val}(x, z.1), z.2) && : e \to A\,, \end{aligned}$$

and finally

$$\bigcup x \quad =_{\mathrm{def}} \quad \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\, A\,.$$

Let $(x : A)\ \phi : \mathsf{prop}$, $a : \mathsf{Sub}(\mathsf{Sub}\, A)$ and assume that $c = \bigcup a : \mathsf{Sub}\, A$. We have

$$\begin{aligned} \mathsf{el}\, c &= (\Sigma y : \mathsf{el}\, a)\, \mathsf{el}(\mathsf{val}(a, y)) && : \mathsf{type}\,, \\ (z : \mathsf{el}\, c) \quad \mathsf{dom}(c, z) &\equiv \mathsf{dom}(a, z.1) \wedge \mathsf{dom}(\mathsf{val}(a, z.1), z.2)\,, \\ (z : \mathsf{el}\, c) \quad \mathsf{val}(c, z) &= \mathsf{val}(\mathsf{val}(a, z.1), z.2) && : A\,, \end{aligned}$$

by the definition of $\bigcup a$. By the computation rules for Σ-types we obtain

$$(y : \mathsf{el}\, a, x : \mathsf{el}(\mathsf{val}(a, y))) \quad \mathsf{val}(c, \mathsf{pair}(x, y)) \quad = \quad \mathsf{val}(\mathsf{val}(a, y), x) \quad : A\,.$$

We now show that

$$(\forall x \in c)\phi \equiv (\forall y \in a)(\forall x \in y)\phi\,.$$

Define

$$\begin{aligned} (y : \mathsf{el}\, a, x : \mathsf{el}(\mathsf{val}(a, y))) \quad \theta &=_{\mathrm{def}} \mathsf{dom}(a, y) && : \mathsf{prop}\,, \\ (y : \mathsf{el}\, a, x : \mathsf{el}(\mathsf{val}(a, y))) \quad \eta &=_{\mathrm{def}} \mathsf{dom}(\mathsf{val}(a, y), x) && : \mathsf{prop}\,. \end{aligned}$$

To prove the desired equivalence it is convenient to consider

$$(\forall z : \mathsf{el}\, c)\, \mathsf{dom}(c, z) \supset \phi[\mathsf{val}(c, z)/x] \tag{8.3}$$

and

$$(\forall y : \mathsf{el}\, a)(\forall x : \mathsf{el}(\mathsf{val}(a, y)))\theta \wedge \eta \supset \phi[\mathsf{val}(c, \mathsf{pair}(x, y))/x]\,. \tag{8.4}$$

We claim that if (8.3) $\equiv$ (8.4) then

$$(\forall x \in c)\phi \equiv (8.3) \equiv (8.4) \equiv (\forall y \in a)(\forall x \in y)\phi\,,$$

which gives the desired result. Assuming that (8.3) $\equiv$ (8.4), we only need to prove the first and the third equivalence: the first one follows by simply unfolding the definitions, while the third one is a consequence of predicate logic rules.

Let us therefore complete the proof by showing that (8.3) $\equiv$ (8.4). First observe that (8.3) implies (8.4) by the $\forall$-elimination rule, and then note that (8.4) implies (8.3) by the Σ-induction rule. The proof of the claim involving the existential quantifier follow the same pattern of reasoning and therefore we omit the details. □

Proposition 8.5 (Small separation). *For* $A : \mathsf{type}$*,* $a : \mathsf{Sub}\, A$ *and* $(x : A)\ p : \mathsf{P}$*, we can define*

$$\{x \in a \mid p\} : \mathsf{Sub}\, A$$

such that if $(x : A)\ \phi : \mathsf{prop}$ *and* $c = \{x \in a \mid p\} : \mathsf{Sub}\, A$ *then the judgements*

$$\begin{aligned} (\forall x \in c)\phi &\equiv (\forall x \in a)(\tau\, p \supset \phi)\,, \\ (\exists x \in c)\phi &\equiv (\exists x \in a)(\tau\, p \wedge \phi)\,, \end{aligned}$$

are derivable.

Proof. For $a : \mathsf{Sub}\, A$ and $(x : A)\ p : \mathsf{P}$ define

$$\begin{aligned} e &=_{\mathrm{def}} \dot{\mathsf{el}}\, a && : \mathsf{U}\,, \\ d &=_{\mathrm{def}} (\lambda x : e)\, \dot{\mathsf{dom}}(a, x) \mathbin{\dot\wedge} p[\mathsf{val}(a,x)/x] && : e \to \mathsf{P}\,, \\ f &=_{\mathrm{def}} (\lambda x : e)\, \mathsf{val}(a, x) && : e \to A\,, \end{aligned}$$

and finally

$$\{x \in a \mid p\ \} \quad =_{\mathrm{def}} \quad \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\, A\,.$$

Let $(x : A)\ \phi : \mathsf{prop}$ and assume that $c = \{x \in a \mid p\} : \mathsf{Sub}\, A$. We have

$$\begin{aligned} & \mathsf{el}\, c &&= \mathsf{el}\, a && : \mathsf{type}\,, \\ (x : \mathsf{el}\, c) \quad & \mathsf{dom}(c, x) &&\equiv \mathsf{dom}(a, x) \wedge \tau\, p[\mathsf{val}(a,x)/x]\,, && \\ (x : \mathsf{el}\, c) \quad & \mathsf{val}(c, x) &&= \mathsf{val}(a, x) && : A\,, \end{aligned}$$

by the definition of $\{x \in a \mid p\}$. The proof is now straightforward:

$$\begin{aligned}(\forall x \in c)\phi \quad &\equiv \quad (\forall x : \mathsf{el}\, a)\big(\big(\,\mathsf{dom}(a,x) \wedge \tau\, p[\mathsf{val}(a,x)/x]\big) \supset \phi[\mathsf{val}(a,x)/x]\big) \\ &\equiv \quad (\forall x : \mathsf{el}\, a)\, \mathsf{dom}(a,x) \supset (\tau\, p[\mathsf{val}(a,x)/x] \supset \phi[\mathsf{val}(a,x)/x]) \\ &\equiv \quad (\forall x \in a)(\tau\, p \supset \phi)\,.\end{aligned}$$

This gives the desired conclusion. The proof of the statement involving existential quantification is analogous and therefore we omit it. □

We will exploit these series of lemmas to show the validity of the axioms of **CZF** under the type-theoretic interpretation in Section 8.5. We conclude this section with two lemmas that are similar in spirit to the previous ones, but will not be used until Chapter 9. Remarkably, to prove the type-theoretic version of replacement described in the next lemma, we do not need any of the collection rules.

Proposition 8.6 (Replacement). *For* $A, B : \mathsf{type}$*,* $a : \mathsf{Sub}\, A$ *and* $(x : A)\ b : B$*, we can define*

$$\{b \mid x \in a\} : \mathsf{Sub}\, B$$

such that if $(y : B)\ \phi : \mathsf{prop}$ *and* $c = \{b \mid x \in a\} : \mathsf{Sub}\, B$ *then the judgements*

$$(\forall x \in a)\phi[b/y] \equiv (\forall y \in c)\phi\,,$$

$$(\exists x \in a)\phi[b/y] \equiv (\exists y \in c)\phi\,,$$

are derivable.

Proof. First of all, observe that we can assume that x is not a free variable in ϕ. Define

$$\begin{aligned} e \quad &=_{\text{def}} \quad \dot{\mathsf{el}}\, a && : \mathsf{U}\,, \\ d \quad &=_{\text{def}} \quad (\lambda x : \mathsf{el}\, a)\, \dot{\mathsf{dom}}(a,x) && : e \to \mathsf{P} \\ f \quad &=_{\text{def}} \quad (\lambda x : \mathsf{el}\, a) b[\mathsf{val}(a,x)/x] && : e \to B\,, \end{aligned}$$

and finally

$$\{b \mid x \in a\} \quad =_{\text{def}} \quad \mathsf{pair}(e, \mathsf{pair}(d,f)) \quad : \mathsf{Sub}\, B\,.$$

Let $(y : B)\ \phi : \mathsf{prop}$ and assume that $c = \{b \mid x \in a\} : \mathsf{Sub}\, B$. We have

$$\begin{array}{rrcll} & \mathsf{el}\, c & = & \mathsf{el}\, a & : \mathsf{type}\,, \\ (x : \mathsf{el}\, c) & \mathsf{dom}(c, x) & \equiv & \mathsf{dom}(a, x)\,, & \\ (x : \mathsf{el}\, c) & \mathsf{val}(c, x) & = & b[\mathsf{val}(a, x)/x] & : B\,, \end{array}$$

by the definition of $\{b \mid x \in a\}$. The remainder of the proof consists in unfolding definitions and substitutions:

$$\begin{array}{rcl} (\forall y \in c)\phi & \equiv & (\forall y : \mathsf{el}\, a)(\mathsf{dom}(a, y) \supset \phi[b[\mathsf{val}(a, y)/x]/y] \\ & \equiv & (\forall x \in a)\phi[b/y]\,, \end{array}$$

where the last step follows because we assumed that x is not a free variable in ϕ. The proof of the statement involving existential quantification is similar and we leave it to the reader. □

We conclude this series of propositions by transferring in logic-enriched type theories a familiar fact of constructive set theories: the correspondence between the class of subsets of a singleton set and restricted sentences, as discussed in Subsection 2.3.3. In logic-enriched type theories the role of subsets of a singleton set is played by the type of small subclasses of the type $\mathsf{1}$ and the role of restricted sentences is played by elements of P.

Proposition 8.7 (Restricted truth values). *For $x : \mathsf{P}$ we can define*

$$\mathsf{ext}\, x : \mathsf{Sub}\, \mathsf{1}$$

such that if $\phi : \mathsf{prop}$, $p : \mathsf{P}$ and $a = \mathsf{ext}\, p : \mathsf{Sub}\, \mathsf{1}$ then the judgements

$$\begin{array}{rcl} \tau\, p & \equiv & (\exists_{-} \in a)\top\,, \\ \tau\, p \supset \phi & \equiv & (\forall_{-} \in a)\phi\,, \end{array}$$

are derivable.

Proof. For $x : \mathsf{P}$ define

$$\begin{array}{rcll} e & =_{\mathrm{def}} & \dot{\mathsf{1}} & : \mathsf{U}\,, \\ d & =_{\mathrm{def}} & (\lambda_{-} : e)x & : e \to \mathsf{P}\,, \\ f & =_{\mathrm{def}} & (\lambda y : e)y & : e \to \mathsf{1}\,, \end{array}$$

and finally

$$\mathsf{ext}\, x \quad =_{\mathrm{def}} \quad \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\, 1\,.$$

The conclusion is a straightforward consequence of the definition of d. We trust the reader capable of filling in the details. □

Before moving on to prove the validity of the axioms of **CZF** we need to have a type to interpret the sets. We define such a type in the next section, and the proceed in Section 8.4 to define the generalised type-theoretic interpretation.

8.3 The type of iterative small classes

We now show that **CZF** can be interpreted in **ML**(**CZF**). Define the type V of **iterative small classes** as

$$\mathsf{V} =_{\mathrm{def}} \big(Wy : (\Sigma x : \mathsf{U})(x \to \mathsf{P})\big) y.1\,.$$

Some intuition. For $e : \mathsf{U}$, $d : \mathsf{T}\, e \to \mathsf{P}$ and $f : \mathsf{T}\, e \to \mathsf{V}$, we may think of a canonical element

$$\mathsf{sup}(\mathsf{pair}(e, d), f) : \mathsf{V}$$

as the ‘set’ of all $\mathsf{app}(f, x) : \mathsf{V}$, with $x : \mathsf{T}\, e$ such that $\tau(\mathsf{app}(d, x))$.

The next lemma illustrates the correspondence between elements of V and elements of $\mathsf{Sub}\,\mathsf{V}$.

Lemma 8.8. *We can define*

$$\begin{aligned} (x : \mathsf{Sub}\,\mathsf{V}) \quad \mathsf{set}\, x \quad &: \quad \mathsf{V}\,, \\ (y : \mathsf{V}) \quad \mathsf{sub}\, y \quad &: \quad \mathsf{Sub}\,\mathsf{V}\,, \end{aligned}$$

such that if $e : \mathsf{U}$*,* $d : e \to \mathsf{P}$ *and* $f : e \to \mathsf{V}$ *then the judgements*

$$\begin{aligned} \mathsf{set}(\mathsf{pair}(e, \mathsf{pair}(d, f))) \quad &= \quad \mathsf{sup}(\mathsf{pair}(e, d), f) \quad : \mathsf{V}\,, \\ \mathsf{sub}(\mathsf{sup}(\mathsf{pair}(e, d), f)) \quad &= \quad \mathsf{pair}(e, \mathsf{pair}(d, f)) \quad : \mathsf{Sub}\,\mathsf{V}\,, \end{aligned}$$

are derivable.

Proof. Define

$$(x : \mathsf{Sub}\,\mathsf{V})\ \ \mathsf{set}\,x =_{\mathrm{def}} \mathsf{sup}(\mathsf{pair}(x.1, x.2.1), x.2.2) : \mathsf{V}$$

and

$$(x : \mathsf{V})\ \ \mathsf{sub}\,x =_{\mathrm{def}} \mathsf{r_W}((u, v)\,\mathsf{pair}(u.1, \mathsf{pair}(u.2, v), x) : \mathsf{Sub}\,\mathsf{V}\,.$$

The desired conclusions can be proved using the Σ-computation rule and the W^{-}-computation rule, respectively. □

Corollary 8.9. *The introduction rule for the type* V *is derivable from the rule*

$$\frac{a : \mathsf{Sub}\,\mathsf{V}}{\mathsf{set}\,a : \mathsf{V}}$$

Proof. The claim is a consequence of Lemma 8.8. □

Proposition 8.10. *For* $(x : \mathsf{V})\ \phi : \mathsf{prop}$ *and* $(y : \mathsf{Sub}\,\mathsf{V})\ \psi : \mathsf{prop}$ *the judgements*

$$(\forall x : \mathsf{V})\ \phi \equiv (\forall y : \mathsf{Sub}\,\mathsf{V})\ \phi[\mathsf{set}\,y/x]\,,$$

$$(\exists x : \mathsf{V})\ \phi \equiv (\exists y : \mathsf{Sub}\,\mathsf{V})\ \phi[\mathsf{set}\,y/x]\,,$$

are derivable.

Proof. To prove the judgement

$$(\forall x : \mathsf{V})\phi \equiv (\forall y : \mathsf{Sub}\,\mathsf{V})\phi[\mathsf{set}\,y/x]\,,$$

we consider the proposition

$$(\forall x : \mathsf{U})(\forall y : \mathsf{T}(x) \to \mathsf{P})(\forall z : \mathsf{T}(x) \to \mathsf{V})\phi[\mathsf{sup}(\mathsf{pair}(x, y), z)/x]\,. \tag{8.5}$$

We indeed show that

$$(\forall x : \mathsf{V})\phi \equiv (8.5) \equiv (\forall y : \mathsf{Sub}\,\mathsf{V})\phi[\mathsf{set}\,y/x]\,. \tag{8.6}$$

The first equivalence in (8.6) can be proved as follows: the 'left-to-right' implication is proved with the ∀-elimination rule, while the 'right-to-left' is proved by W^{-}-induction. The second equivalence in (8.6) can instead be obtained as we describe now. The 'left-to-right' implication is a consequence of the Σ-induction rule, while the 'right-to-left' implication follows by the ∀-elimination rule and Lemma 8.8. □

We will now define an extensional equality on the type V of **ML(CZF)** that will be used to interpret the set-theoretic equality of **CZF**. An inspection of the original type-theoretic interpretation reveals that Π-types are used to define an extensional equality on the type of iterative sets [3, Theorem 2.3]. Since in **ML(CZF)** has only restricted forms of Π-types, we rather use the double W^--elimination rule of **ML(CZF)** given in Section B.3.

Before turning to the extensional equality, however, we combine the convenient notation of 'restricted quantifiers' introduced on page 7.1 with the correspondence between elements of V and $\mathsf{Sub}\,\mathsf{V}$ of Lemma 8.8. The combination of these definitions allows us to get a 'restricted quantification' on elements of V. For $x : \mathsf{V}$, $(y : \mathsf{V})\ p : \mathsf{P}$, and $(y : \mathsf{V})\ \phi : \mathsf{prop}$ define

$$
\begin{aligned}
(\dot{\nabla} y \in x)p &=_{\text{def}} (\dot{\nabla} y \in \mathsf{sub}\,x)p : \mathsf{P}\,, \\
(\nabla y \in x)\phi &=_{\text{def}} (\nabla y \in \mathsf{sub}\,x)\phi : \mathsf{prop}\,,
\end{aligned}
$$

where ∇ is either $\forall$ or $\exists$. These definitions will be implicitly used throughout the remainder of this chapter.

Proposition 8.11. *We can define*

$$(x, y : \mathsf{V})\ x \dot{\approx} y : \mathsf{P}$$

such that if $a, b : \mathsf{V}$ *then the judgement*

$$a \approx b \equiv (\forall x \in a)(\exists y \in b)x \approx y \wedge (\forall y \in b)(\exists x \in a)x \approx y\,,$$

where $a \approx b =_{\text{def}} \tau(a \dot{\approx} b)$, *is derivable.*

Proof. For $u_1 : \mathsf{S}$, $v_1 : u_1.1 \to \mathsf{V}$, $u_2 : \mathsf{S}$, $v_2 : u_2.1 \to \mathsf{V}$, $w : u_1.1 \to (u_2.1 \to \mathsf{P})$ define $d_1, d_2 : \mathsf{P}$ such that if $\phi_1 \equiv \tau d_1$ and $\phi_2 \equiv \tau d_2$ then

$$
\begin{aligned}
\phi_1 &\equiv (\forall x : u_1.1)\,\mathsf{app}(u_1.2, x) \to (\exists y : u_2.1)(\mathsf{app}(u_2.2, y) \wedge \mathsf{app}(\mathsf{app}(w, x), y)\,, \\
\phi_2 &\equiv (\forall y : u_2.1)\,\mathsf{app}(u_2.2, y) \to (\exists x : u_1.1)(\mathsf{app}(u_1.2, x) \wedge \mathsf{app}(\mathsf{app}(w, x), y)\,,
\end{aligned}
$$

are derivable, and define $d =_{\text{def}} d_1 \wedge d_2 : \mathsf{P}$. We can now apply the double W^--elimination rule of **ML(CZF)** of Section B.3 and define

$$(x, y : \mathsf{V})\ x \dot{\approx} y =_{\text{def}} \mathsf{r_{2W}}(x, y, (u_1, v_1, u_2, v_2, w)d) : \mathsf{P}\,.$$

The W^--induction and the double W^--computation rule now lead us to conclude the proof. □

Proposition 8.11 concludes our preparations for the definition of the generalised type-theoretic interpretations of **CZF**, that is presented in the next section.

8.4 Sets-as-trees

We assume that **CZF** is formulated in the language $\mathcal{L}$, that has no constants, and that the symbols for variables of the language $\mathcal{L}$ coincide with ones of type V. Recall from Section 2.1 that the language $\mathcal{L}$ has primitive restricted quantifiers, but no primitive membership relation. We define two interpretations: a first one applies to arbitrary formulas, and another interpretation applies only to restricted formulas. Table 8.1 on page 165 and Table 8.2 on page 165 contain the definitions of the interpretations of arbitrary and restricted formulas, respectively.

The next lemma shows that arbitrary and restricted formulas of $\mathcal{L}$ correspond to arbitrary and small propositions of **ML**(**CZF**).

Lemma 8.12. *If ϕ is a formula of $\mathcal{L}$ with $\mathrm{FV}\phi = \vec{x}$ and θ is a restricted formula of $\mathcal{L}$ with $\mathrm{FV}\theta = \vec{x}$ then the judgements*

$$(\vec{x} : \mathsf{V}) \quad [\![\, \phi \,]\!] : \mathsf{prop}\,,$$

$$(\vec{x} : \mathsf{V}) \quad (\!|\, \theta \,|\!) : \mathsf{P}\,,$$

$$(\vec{x} : \mathsf{V}) \quad \tau(\!|\, \theta \,|\!) \equiv [\![\, \theta \,]\!]$$

are derivable.

Proof. Direct consequence of the definition of the interpretations. □

Definition 8.13. We say that the interpretation of a formula ϕ of $\mathcal{L}$ with $\mathrm{FV}\ \phi = \vec{x}$ is **sound** if the judgement

$$(\vec{x} : \mathsf{V}) \quad [\![\phi]\!]$$

is derivable. We say that the generalized type theoretic interpretation of a constructive set theory is **sound** if the interpretation of all its axioms and of all the instances of its axiom schemes is valid. ◊

8.5 Soundness

In this section we prove the following theorem.

Theorem 8.14. *The generalised type theoretic interpretation of* **CZF** *is sound.*

Table 8.1: Sets-as-trees interpretation for arbitrary formulas.

$$
\begin{aligned}
[\![x = y]\!] &=_{\mathrm{def}} x \approx y\,,\\
[\![\phi_1 \wedge \phi_2]\!] &=_{\mathrm{def}} [\![\phi_1]\!] \wedge [\![\phi_2]\!]\,,\\
[\![\phi_1 \vee \phi_2]\!] &=_{\mathrm{def}} [\![\phi_1]\!] \vee [\![\phi_2]\!]\,,\\
[\![\phi_1 \rightarrow \phi_2]\!] &=_{\mathrm{def}} [\![\phi_1]\!] \supset [\![\phi_2]\!]\,,\\
[\![(\forall x \in y)\, \phi_0]\!] &=_{\mathrm{def}} (\forall x \in y)\, [\![\phi_0]\!]\,,\\
[\![(\exists x \in y)\, \phi_0]\!] &=_{\mathrm{def}} (\exists x \in y)\, [\![\phi_0]\!]\,,\\
[\![(\forall x)\, \phi_0]\!] &=_{\mathrm{def}} (\forall x : \mathsf{V})\, [\![\phi_0]\!]\,,\\
[\![(\exists x)\, \phi_0]\!] &=_{\mathrm{def}} (\exists x : \mathsf{V})\, [\![\phi_0]\!]\,.
\end{aligned}
$$

Table 8.2: Sets-as-trees interpretation for restricted formulas.

$$
\begin{aligned}
(\!|x = y|\!) &=_{\mathrm{def}} x \mathrel{\dot{\approx}} y\,,\\
(\!|\theta_1 \wedge \theta_2|\!) &=_{\mathrm{def}} (\!|\theta_1|\!) \mathbin{\dot{\wedge}} (\!|\theta_2|\!)\,,\\
(\!|\theta \vee \theta_2|\!) &=_{\mathrm{def}} (\!|\theta_1|\!) \mathbin{\dot{\vee}} (\!|\theta_2|\!)\,,\\
(\!|\theta_1 \rightarrow \theta_2|\!) &=_{\mathrm{def}} (\!|\theta_1|\!) \mathbin{\dot{\supset}} (\!|\theta_2|\!)\,,\\
(\!|(\forall x \in y)\, \theta_0|\!) &=_{\mathrm{def}} (\dot{\forall} x \in y)\, (\!|\theta_0|\!)\,,\\
(\!|(\exists x \in y)\, \theta_0|\!) &=_{\mathrm{def}} (\dot{\exists} x \in y)\, (\!|\theta_0|\!)\,.
\end{aligned}
$$

The proof of this theorem will follow by the series of lemmas 8.16, 8.17 and 8.18. We start with a simple lemma.

Lemma 8.15. *Let* $(x : \mathsf{V})\ \phi : \mathsf{prop}$ *and* $a, b : \mathsf{V}$*. Then the judgement*

$$[\![\phi[a/x]]\!] \wedge [\![a = b]\!] \supset [\![\phi[b/x]]\!]$$

is derivable.

Proof. By structural induction on ϕ. ☐

Given lemma 8.15 it is straightforward to observe that all the logical axioms for **CZF**, and in particular the ones regarding restricted quantifiers, are sound. The next lemma takes care of the structural axioms of **CZF**.

Lemma 8.16. *The interpretation of the Extensionality and Set Induction is sound.*

Proof. Soundness of Extensionality is a consequence of proposition 8.11. Soundness of Set Induction follows by the W^{-}-induction rule. □

We now pick the fruits of the results proved in Section 8.2. Proofs of the soundness of the interpretation of the set existence axioms of **CZF** are simple consequences of lemmas 8.1 – 8.6. Some work is required only for Infinity.

Lemma 8.17. *The interpretation of Pairing, Union, Infinity and Restricted Separation is sound.*

Proof. Lemma 8.2 helps us to prove validity of Pairing. Let $a, b : \mathsf{V}$ and define $c =_{\text{def}} \mathsf{set}\{a, b\} : \mathsf{V}$. Observe that

$$\begin{aligned} (\forall x \in c)(x \approx a \vee x \approx b) &\equiv (\forall x \in \{a, b\})(x \approx a \vee x \approx b) \\ &\equiv (a \approx a \vee a \approx b) \wedge (b \approx a \vee b \approx b) \end{aligned}$$

by the definition of 'restricted quantification' over $c : \mathsf{V}$, the definition of $\{a, b\}$ and Lemma 8.2. Similarly we get

$$\begin{aligned} a \in c \wedge b \in c &\equiv (\exists x \in \{a, b\})(x \approx a) \wedge (\exists x \in \{a, b\})(x \approx b) \\ &\equiv (a \approx a \vee a \approx b) \wedge (b \approx a \vee b \approx b) \end{aligned}$$

as wanted. Soundness of Union follows in a similar way from Lemma 8.4. For Infinity, recall Lemma 8.1 and Corollary 8.3. First of all define

$$(x : \mathsf{N}, y : \mathsf{V}) \quad s \quad =_{\text{def}} \quad \mathsf{set}\{y\} \quad : \mathsf{V}$$

Now define

$$\begin{aligned} e &=_{\text{def}} \dot{\mathsf{N}} &&: \mathsf{U}\,, \\ d &=_{\text{def}} (\lambda z : e)\dot{\top} &&: e \to \mathsf{P}\,, \\ f &=_{\text{def}} (\lambda z : e)\, \mathsf{r_N}(\mathsf{set}(\emptyset), (x, y)s, z) &&: e \to \mathsf{V}\,, \end{aligned}$$

where we used the computation rule for the type N given in Subsection B.2.1, and finally

$$c \quad =_{\text{def}} \quad \mathsf{pair}(\mathsf{pair}(e, d), f) \quad : \mathsf{Sub}\,\mathsf{V}\,.$$

It is now immediate to see that $\mathsf{set}\, c : \mathsf{V}$ can be used to show the soundness of Infinity. Validity of Restricted Separation follows from Lemma 8.5. We leave the details to the reader. □

The next lemma should not be a surprise.

Lemma 8.18. *The interpretation of Strong Collection and Subset Collection is sound.*

Proof. Straightforward consequence of the collection rules of **ML**(**CZF**) given in Table 7.4. □

We have therefore proved Theorem 8.14, stating that **CZF** is interpretable in **ML**(**CZF**).

8.6 Types-as-classes

In this section we recall a result obtained in [6] that implies the mutual interpretability of **CZF** and **ML**(**CZF**). From now on we work informally in **CZF**. We say that ξ is a **variable assignment** if $\xi(x)$ is a set for each variable x. Recall also from Subsection 2.1.1 that ϕ is a set theoretical sentence if it has no free variables.

The terminology used in relation to the raw syntax for logic-enriched type theories in Subsection 3.1.2 and Subsection 7.2.1 may help the reader to understand why we will say that a set is a **0-class**, a class is a **1-class** and a set theoretical sentence is a **2-class**. Also, for $n \geq 0$ and $\epsilon = 0, 1, 2$, we say that a definable operator F is an n^ϵ**-class** if it assigns an ϵ-class $F(a_1, \ldots, a_n)$ to each n-tuple $a_1, \ldots, a_n$ of sets.

Definition 8.19. We say that an assignment $[\![M]\!]_\xi$ to each expression M and each variable assignment ξ of the raw syntax is a **types-as-classes interpretation** if $[\![a]\!]_\xi$ is a set for each term a, $[\![A]\!]_\xi$ is a class for each type expression A and $[\![\phi]\!]_\xi$ is a set theoretical sentence for each proposition expression ϕ. ◊

Given a set theoretical interpretation $\mathcal{F}_\kappa$ of each symbol κ of the signature, it is possible to define a types-as-classes interpretation by structural induction on the way expressions are built up. Of course, the set theoretical interpretation of the symbols of the signature has to be such that the second clause of the following definition by structural induction makes sense:

$$\begin{aligned} [\![x]\!]_\xi &=_{\text{def}} \xi(x)\,, \\ [\![\kappa((\vec{x}_1)M_1, \ldots, (\vec{x}_k)M_k)]\!] &=_{\text{def}} \mathcal{F}_\kappa(F_1, \ldots, F_k)\,, \end{aligned}$$

where κ is a constant symbol of the signature of arity $(n_1^{\epsilon_1} \cdots n_k^{\epsilon_k})^\epsilon$ and, for F_i is the $n_i^{\epsilon_i}$-class such that for $i = 1, \ldots, k$

$$F_i(\vec{a}_i) = [\![M_i]\!]_{\xi(\vec{a}_i/\vec{x}_i)}$$

for all n_i-tuples $\vec{a}_i$ of sets. When κ has arity $(n_1^{\epsilon_1} \cdots n_k^{\epsilon_k})^{\epsilon}$ we require that $\mathcal{F}_\kappa$ is a set operator of that arity. This means that whenever F_i is an $n_i^{\epsilon_i}$-class, for $i = 1, \dots, k$ then $\mathcal{F}_\kappa(F_1, \dots, F_k)$ should be an ϵ-class obtained 'uniformly' from $F_1, \dots, F_k$. Given a types-as-classes interpretation as above we can define the notions in Table 8.3

Table 8.3: Soundness of the types-as-classes interpretation.

$$\begin{aligned}
\xi \models (x_1 : A_1, \dots, x_n : A_n) &=_{\text{def}} \xi(x_1) \in [\![A_1]\!]_\xi \wedge \cdots \wedge \xi(x_n) \in [\![A_n]\!]_\xi \,, \\
\xi \models A : \mathsf{type} &=_{\text{def}} \top \,, \\
\xi \models A_1 = A_2 : \mathsf{type} &=_{\text{def}} [\![A_1]\!]_\xi = [\![A_2]\!]_\xi \,, \\
\xi \models a : A &=_{\text{def}} [\![a]\!]_\xi \in [\![A]\!]_\xi \,, \\
\xi \models a_1 = a_2 : A &=_{\text{def}} [\![a_1]\!]_\xi = [\![a_2]\!]_\xi \,, \\
\xi \models \phi : \mathsf{prop} &=_{\text{def}} \top \,, \\
\xi \models \phi_1 \dots, \phi_m \Rightarrow \phi &=_{\text{def}} [\![\phi_1]\!]_\xi \wedge \cdots \wedge [\![\phi_m]\!]_\xi \to [\![\phi]\!]_\xi \,.
\end{aligned}$$

Definition 8.20. We say that the types-as-classes interpretation of a raw judgement $(\Gamma)\ \mathcal{B}$ is **sound** if $\xi \models \Gamma$ implies $\xi \models \mathcal{B}$ for every variable assignment ξ. The types-as-classes interpretation of a type theoretic rule is **sound** if whenever the premisses of an instance of the rule are valid then so is the conclusion. ♢

Theorem 8.21 (Aczel). **ML(CZF)** *has a sound the types-as-classes interpretation.*

Proof. See [6] and also [5] for more details. □

Corollary 8.22. *The systems* **CZF** *and* **ML(CZF)** *are mutually interpretable.*

Proof. The claim follows from Theorem 8.14 and Theorem 8.21. □

Chapter 9

Reinterpreting logic

9.1 Introduction

The aim of this chapter is to show how the logic-enriched type theory **ML**(**CZF**) accommodates reinterpretations of logic. We start by considering a propositions-as-types translation into a pure type theory, and observe how collection rules are consequences of the type-theoretic axiom of choice.

We then consider reinterpretations defined by considering an operator j on the proposition universe that satisfies type-theoretic versions of the properties of a nucleus on a set-generated frame, as introduced in Definition 4.9 on page 55, or of a Lawvere-Tierney operator, as described in Definition 6.18 on page 128. We call the reinterpretations determined by such an operators **j-interpretations**. These can be defined following the definition of the syntactic translation determined by a Lawvere-Tierney operator, as in Section 6.7.

Collection rules play a prominent role in j-interpretations. First of all, the Strong Collection rule allows to show that an extension of j from small to arbitrary propositions inherits the properties of j. It is actually using this fact that we can set up j-interpretations of logic. Furthermore, the Strong Collection rule is preserved by the j-interpretations, in intuitive analogy with the development of the second step of sheaf interpretations for CST developed in Chapter 6. As one might expect, the Subset Collection rule does not seem instead to be preserved by arbitrary j-interpretations. The intuitive analogy with set theory leads us to formulate a suitable condition on j that allows us to show that the Subset Collection rule is also preserved. We end the chapter with a type-theoretic version of the double-negation interpretation.

9.2 Propositions-as-types translation

In this section we show how the logic-enriched type theory $\mathbf{ML}(\mathbf{CZF})$ has a translation into the pure type theory $\mathbf{ML_1} + \mathbf{W}^-$, that is obtained by adding W^--rules to the type theory $\mathbf{ML_1}$. Note that the pure type theory we consider has Π-types and not just Π^--types. This is necessary to define the translation of universal quantification on types that are not necessarily small. To formulate the propositions-as-types translation, it is convenient to extend the pure type theory $\mathbf{ML_1} + \mathbf{W}^-$ with a new form of judgement

$$(\Gamma)\ B_1, \ldots, B_m \Rrightarrow B\,. \tag{9.1}$$

This judgement is well-formed if all the judgements $(\Gamma)\ B_i : \mathsf{type}$ for $i = 1, \ldots, m$ and $(\Gamma)\ B : \mathsf{type}$ are derivable. This extension is obviously conservative, since the only rule involving this judgement is the following:

$$\frac{(y_1 : B_1, \ldots, y_m : B_m)\ b : B}{B_1, \ldots, B_m \Rrightarrow B\,.}$$

Let us now define some expressions that will be useful in the following. We invite the reader to compare them with the definitions of non-dependent products, binary sums, and non-dependent functions types given in Subsection 3.3.1.

$$\begin{array}{llll}
(x_1, x_2 : \mathsf{U}) & x_1 \dot{\times} x_2 & =_{\mathrm{def}} (\dot{\Sigma}_- : x_1) x_2 & : \mathsf{U}\,, \\
(x_1, x_2 : \mathsf{U}) & x_1 \dot{+} x_2 & =_{\mathrm{def}} (\dot{\Sigma} z : \dot{2})\, \mathsf{r}_2(x_1, x_2, z) & : \mathsf{U}\,, \\
(x_1, x_2 : \mathsf{U}) & x_1 \dot{\rightarrow} x_2 & =_{\mathrm{def}} (\dot{\Pi}_- : x_1) x_2 & : \mathsf{U}\,.
\end{array}$$

While the propositions-as-types translation for a logic-enriched type theory without a proposition universe is rather straightforward, some attention is needed to extend this interpretation also to the expressions associated to a proposition universe, in order to respect the distinction between small propositions and representatives for them. We will therefore define the propositions-as-types translation as the composition of two translations, that we now define separately.

The first translation, that we indicate with $\langle \cdot \rangle$, maps raw expressions regarding the type P into raw expressions regarding the type U, and it is defined in Table 9.1. The second translation, that we indicate with $[\![\cdot]\!]$, is the straightforward propositions-as-types translation for predicate logic, and it is defined in Table 9.2.

As the reader may have already noticed, the definition of the $[\![\cdot]\!]$-translation is not complete, as we still need to define the translation of expressions of the form $\tau\, p$ for $p : \mathsf{P}$. This is not a problem, because it is exactly for these expressions that we introduced

Table 9.1: Propositions-as-types translation for a proposition universe.

$$
\begin{aligned}
\langle \mathsf{P} \rangle &=_{\text{def}} \mathsf{U}\,, \\
\langle \dot{\bot} \rangle &=_{\text{def}} \dot{\mathsf{0}}\,, \\
\langle \dot{\top} \rangle &=_{\text{def}} \dot{\mathsf{1}}\,, \\
\langle p_1 \mathbin{\dot{\wedge}} p_2 \rangle &=_{\text{def}} \langle p_1 \rangle \mathbin{\dot{\times}} \langle p_2 \rangle\,, \\
\langle p_1 \mathbin{\dot{\vee}} p_2 \rangle &=_{\text{def}} \langle p_1 \rangle \mathbin{\dot{+}} \langle p_2 \rangle\,, \\
\langle p_1 \mathbin{\dot{\supset}} p_2 \rangle &=_{\text{def}} \langle p_1 \rangle \mathbin{\dot{\rightarrow}} \langle p_2 \rangle\,, \\
\langle (\dot{\forall} x : a) p_0 \rangle &=_{\text{def}} (\dot{\Pi} x : a) \langle p_0 \rangle\,, \\
\langle (\dot{\exists} x : a) p_0 \rangle &=_{\text{def}} (\dot{\Sigma} x : a) \langle p_0 \rangle\,.
\end{aligned}
$$

Table 9.2: Propositions-as-types translations for predicate logic.

$$
\begin{aligned}
[\![\bot]\!] &=_{\text{def}} \mathsf{0}\,, \\
[\![\top]\!] &=_{\text{def}} \mathsf{1}\,, \\
[\![\phi_1 \wedge \phi_2]\!] &=_{\text{def}} [\![\phi_1]\!] \times [\![\phi_2]\!]\,, \\
[\![\phi_1 \vee \phi_2]\!] &=_{\text{def}} [\![\phi_1]\!] + [\![\phi_2]\!]\,, \\
[\![\phi_1 \supset \phi_2]\!] &=_{\text{def}} [\![\phi_1]\!] \rightarrow [\![\phi_2]\!]\,, \\
[\![(\forall x : A) \phi_0]\!] &=_{\text{def}} (\Pi x : A) [\![\phi_0]\!]\,, \\
[\![(\exists x : A) \phi_0]\!] &=_{\text{def}} (\Sigma x : A) [\![\phi_0]\!]\,.
\end{aligned}
$$

the first translation: we indeed define

$$[\![\tau\, p]\!] =_{\text{def}} \mathsf{T}\, \langle p \rangle\,.$$

Observe that if $p : \mathsf{P}$, then $\langle p \rangle : \mathsf{U}$ and hence $[\![\tau\, p]\!] : \mathsf{type}$.

The two translations can then be extended to deal with other expressions apart from the ones regarding logic, and also on judgements. We extend $\langle \cdot \rangle$ by letting it be the identity on all the other expressions of **ML**(**CZF**). For a judgement J of **ML**(**CZF**), define $\langle J \rangle$ as the judgement that is obtained by replacing all the expressions in J according to the $\langle \cdot \rangle$-translation. The $[\![\cdot]\!]$-translation leaves unchanged all the raw judgements of **ML**(**CZF**) that are judgements of a pure type theory. For the other judgements define

$$[\![(\Gamma)\ \phi : \mathsf{prop}]\!] \quad =_{\text{def}} \quad (\Gamma)\ [\![\phi]\!] : \mathsf{type}\,,$$

$$[\![(\Gamma)\ \phi_1, \ldots, \phi_m \Rightarrow \phi]\!] \quad =_{\text{def}} \quad (\Gamma)\ [\![\phi_1]\!], \ldots, [\![\phi_m]\!] \Rightarrow [\![\phi]\!]\,,$$

where we made use of the judgement introduced in (9.1) at the beginning of this section. The following result was obtained in [6]. The next result has been obtained in collaboration with Peter Aczel.

Theorem 9.1. *If* $\mathbf{ML}(\mathbf{CZF}) \vdash J$, *then* $\mathbf{ML_1} + \mathbf{W}^- \vdash [\![\langle J \rangle]\!]$.

Proof. We only discuss why the translation of the collection rules is derivable, but let us also point out that the extra W^--elimination rule of Section B.3 is derivable once full Π-rules are assumed, as in $\mathbf{ML_1} + \mathbf{W}^-$. To discuss collection rules, let us define

$$(x : \mathsf{U})\ \ !\,x =_{\text{def}} (\exists_{\_} : \mathsf{T}\,x)\top\,.$$

In (9.2), where $C =_{\text{def}} (\Pi^- x : a)B$, we express a version of the type-theoretic axiom of choice, that can be formulated in **ML**(**CZF**).

$$\frac{a : \mathsf{U} \qquad (x : \mathsf{T}\,a)\ B : \mathsf{type} \qquad (x : \mathsf{T}\,a, y : B)\ \phi : \mathsf{prop}}{(\forall x : \mathsf{T}\,a)(\exists y : B)\phi \Rightarrow (\exists z : C)(\forall x : \mathsf{T}\,a)\phi[\mathsf{app}(z, x)/y]\,.} \tag{9.2}$$

We also need a rule that expresses a correspondence between small types and small propositions.

$$\frac{p : \mathsf{P}}{\tau\, p \equiv\, !\langle p \rangle\,.} \tag{9.3}$$

The key fact to show that the translation of the collection rules is derivable is that both of them follow from the combination of rules (9.2) and (9.3), as direct derivations might show. Furthermore, the propositions-as-types translations of (9.2) and (9.3) are derivable in $\mathbf{ML_1} + \mathbf{W}^-$. In particular, the proof that the translation of (9.2) is derivable follows Martin-Löf's original derivation of the type-theoretic axiom of choice [59, pages 50 – 52]. □

9.3 j-interpretations

Recall from Subsection 7.6.2 that the logic-enriched type theory $\mathbf{ML}(\mathbf{CZF}^-)$ is obtained from **ML**(**CZF**) by omitting the Subset Collection rule. In discussing j-interpretations, it seems appropriate to consider the logic-enriched type theory $\mathbf{ML}(\mathbf{CZF}^-)$ initially, and $\mathbf{ML}(\mathbf{CZF})$ at a later stage. There are two main reasons for doing so. A first reason is that the Strong Collection rule is sufficient to prove the basic properties of j-interpretations. A second reason is that the Strong Collection rule is preserved by the j-interpretation determined by any topology j, while the Subset Collection rule does not seem to be.

The next abbreviation will convenient to state Definition 9.2. For $a, b : \mathsf{P}$ define

$$a \leq b =_{\text{def}} \tau(a) \supset \tau(b) : \mathsf{prop}\,.$$

Definition 9.2. Let j be an explicitly defined operator on P, i.e. assume that there is an explicit definition of the form $(x : \mathsf{P})\ jx =_{\text{def}} e : \mathsf{P}$, where $(x : \mathsf{P})\ e : \mathsf{P}$. We say that j is a **topology** if the judgements

- j is inflationary, i.e. $(\forall x : \mathsf{P})\, x \leq jx\,,$
- j is monotone, i.e. $(\forall x, y : \mathsf{P})\, x \leq y \supset jx \leq jy\,,$
- j is idempotent, i.e. $(\forall x : \mathsf{P})\, j(jx) \leq jx\,,$
- j respects meets, i.e. $(\forall x, y : \mathsf{P})\, jx \mathbin{\dot\wedge} jy \leq j(x \mathbin{\dot\wedge} y)\,,$

are derivable. ◊

From now on we work informally within $\mathbf{ML}(\mathbf{CZF}^-)$ and consider a fixed, arbitrary topology j. The next definition will be very convenient in the following. For $y : \mathsf{P}$ and $\phi : \mathsf{prop}$ define

$$y \leq \phi =_{\text{def}} \tau\, y \supset \phi\,.$$

In the remainder of this section we proceed in analogy with our development of Heyting-valued and sheaf interpretations in Chapter 5 and Chapter 6.

First of all, we introduce the definition of a modality J that extends j to arbitrary propositions. The reader is invited to compare this definitions with the ones given in Section 5.2 and Section 6.7 in the context of Heyting-valued and sheaf interpretations for constructive set theories. For $\phi : \mathsf{prop}$, define

$$J\phi =_{\text{def}} (\exists x : \mathsf{P})(\tau(jx) \wedge x \leq \phi)\,.$$

The next proposition shows that J extends j.

Proposition 9.3. *For $a : \mathsf{P}$, $J(\tau\, a) \equiv \tau(ja)$.*

Proof. The proof consists in a direct calculation. ∎

We will use this definition in the statement and the proof of the next lemma, that we invite the reader to compare with Proposition 2.5.

Lemma 9.4. *Let* $A : \mathsf{type}$*,* $a : \mathsf{Sub}\, A$*,* $\phi : \mathsf{prop}$*,* $(x : A, y : \mathsf{P})\ \psi : \mathsf{prop}$ *and assume that*

$$(\forall x \in a)\big((\exists y : \mathsf{P})((y \leq \phi) \wedge \psi) \wedge (\forall y, z : \mathsf{P})((y \leq z \leq \phi \wedge \psi) \supset \psi[z/y])\big)$$

There is $q : \mathsf{P}$ *such that* $q \leq \phi \wedge (\forall x \in a)\psi[q/y]$*.*

Proof. Define

$$(x : A, y : \mathsf{P})\ \xi =_{\mathrm{def}} y \leq \phi \wedge \psi : \mathsf{prop}\,.$$

By the assumption in the claim and the definition of ξ , we have

$$(\forall x \in a)(\exists y : \mathsf{P})\xi$$

By the Strong Collection rule there is $u : \mathsf{Sub}\,\mathsf{P}$ such that

$$\mathsf{coll}(x \in a, y \in u, \xi)\,. \tag{9.4}$$

Once we define

$$q =_{\mathrm{def}} (\dot{\exists} y \in u)y : \mathsf{P}\,,$$

the definition of ξ, the assumption in the claim and (9.4) imply that

$$q \leq \phi \wedge (\forall x \in a)\psi[q/y]\,.$$

Discharging the assumption of u we get the desired conclusion. □

The next lemma is analogous, for example, to Lemma 5.2.

Lemma 9.5. *Let* $\phi : \mathsf{prop}$*, let* $p : \mathsf{P}$ *and assume that* $p \leq J\phi$*. Then there is* $q : \mathsf{P}$ *such that* $q \leq \phi \wedge p \leq jq$*.*

Proof. Assume that $p \leq J\phi$. Recalling Lemma 8.7, define

$$a =_{\mathrm{def}} \mathsf{ext}\, p : \mathsf{Sub}\, 1$$

and

$$(\_ : 1, y : \mathsf{P})\ \psi =_{\mathrm{def}} \tau(jy) : \mathsf{prop}\,.$$

Lemma 8.7 and the assumption of the claim imply that

$$(\forall x \in a)\big((\exists y : \mathsf{P})(y \leq \phi \wedge \psi) \wedge (\forall y, z : \mathsf{P})((y \leq z \leq \phi \wedge \psi) \supset \psi[z/y])\big)\,,$$

because j is monotone. Lemma 9.4 can now be applied to obtain that there is $q : \mathsf{P}$ such that

$$q \leq \phi \wedge (\forall x \in a)\psi[q/y]\,,$$

and therefore

$$q \leq \phi \wedge p \leq jq$$

by the definitions of a of ψ, and Lemma 8.7. □

The next proposition shows that the properties of j can be lifted to J. Its proof is based on Lemma 9.5 and therefore seems to make essential use of the Strong Collection rule. We invite the reader to compare it with Proposition 5.3.

Proposition 9.6. *For $\phi_1, \phi_2 : \mathsf{prop}$, the following judgements*

(i) $\phi_1 \supset J\phi_1\,,$

(ii) $\phi_1 \supset \phi_2 \Rightarrow J\phi_1 \supset J\phi_2\,,$

(iii) $J(J\phi_1) \supset J\phi_1\,,$

(iv) $J\phi_1 \wedge J\phi_2 \supset J(\phi_1 \wedge \phi_2)\,.$

are derivable.

Proof. For (i), (ii) and (iv) direct calculations suffice. For (iii) use Lemma 9.5 and the fact that j is monotone. □

We now define the j-interpretation of of $\mathbf{ML}(\mathbf{CZF}^-)$ into itself determined by the topology j. This interpretation acts solely on the logic, leaving types unchanged. We define the j-interpretation $\langle\cdot\rangle$ by structural induction on the raw syntax of the type theory, in complete analogy to the definition presented in Table 6.2 on page 130 in the context of sheaf interpretations. Type expressions are left unchanged. The definition of the j-interpretation of formulas is contained in Table 9.3.

All the jugdement bodies that are part of a pure type theory are left unchanged. For the other judgement bodies define:

$$\langle \phi : \mathsf{prop} \rangle \quad =_{\mathrm{def}} \quad \langle \phi \rangle : \mathsf{prop}\,,$$

$$\langle\; \phi_1, \ldots, \phi_n \Rightarrow \phi \;\rangle \quad =_{\mathrm{def}} \quad J\langle \phi_1 \rangle, \cdots, J\langle \phi_n \rangle \Rightarrow J\langle \phi \rangle$$

$$
\begin{aligned}
\langle \top \rangle &=_{\text{def}} \top \\
\langle \bot \rangle &=_{\text{def}} \bot\,, \\
\langle \phi_1 \wedge \phi_2 \rangle &=_{\text{def}} J\langle \phi_1 \rangle \wedge J\langle \phi_2 \rangle\,, \\
\langle \phi_1 \vee \phi_2 \rangle &=_{\text{def}} J\langle \phi_1 \rangle \vee J\langle \phi_2 \rangle\,, \\
\langle \phi_1 \supset \phi_2 \rangle &=_{\text{def}} J\langle \phi_1 \rangle \supset J\langle \phi_2 \rangle\,, \\
\langle (\forall x : A)\phi_0 \rangle &=_{\text{def}} (\forall x : A) J\langle \phi_0 \rangle\,, \\
\langle (\exists x : A)\phi_0 \rangle &=_{\text{def}} (\exists x : A) J\langle \phi_0 \rangle\,, \\
\langle \tau(a) \rangle &=_{\text{def}} \tau(ja)\,.
\end{aligned}
$$

Table 9.3: j-reinterpretation of formulas.

Finally, define the j-interpretation of judgements as follows.

$$\langle (\Gamma)\ \mathcal{B} \rangle =_{\text{def}} (\Gamma)\ \langle \mathcal{B} \rangle.$$

Definition 9.7. We say that the j-interpretation of a rule is **sound** if, assuming the j-interpretation of premisses, the j-interpretation of the conclusion is derivable. $\lozenge$

9.4 Soundness of the j-interpretation

We now show that the j-interpretation of all the rules of $\mathbf{ML}(\mathbf{CZF}^-)$ is sound. We begin with a simple observation.

Proposition 9.8. *The j-interpretation of the predicate logic and induction rules of $\mathbf{ML}(\mathbf{CZF})$ is sound.*

Proof. The result follows by a series of routine calculations. $\square$

The next proposition will be used in the proof of Lemma 9.10.

Proposition 9.9. *For $\phi : \mathsf{prop}$, the judgement*

$$J\phi \equiv (\exists p : \mathsf{Sub}\,1)\big(J(\exists_{\_} \in p)\top \wedge (\forall_{\_} \in p)\phi\big)$$

is derivable.

Proof. The claim is a consequence of the definition of J and of Lemma 8.7. $\square$

The next lemma is crucial to prove that the j-interpretation of the Strong Collection rule of $\mathbf{ML}(\mathbf{CZF}^-)$ is sound.

Lemma 9.10. *For* $A, B : \mathsf{type}$*,* $(x : A, y : B)\ \phi : \mathsf{prop}$*,* $a : \mathsf{Sub}(A)$*, the judgement*

$$(\forall x \in a)\ J(\exists y : B)\phi \Rightarrow (\exists v : \mathsf{Sub}\,B)\big((\forall x \in a)\ J(\exists y \in v)\phi \wedge (\forall y \in v)(\exists x \in a)\phi\big)$$

is derivable.

Proof. Define

$$(p : \mathsf{Sub}\,1)\ \sigma =_{\mathrm{def}} J(\exists\_ \in p)\top : \mathsf{prop}\,,$$

and assume that

$$(\forall x \in a)\ J(\exists y : B)\phi\,.$$

We get

$$(\forall x \in a)(\exists p : \mathsf{Sub}\,1)\ \big(\sigma \wedge (\forall\_ : p)(\exists y : B)\phi\big) \tag{9.5}$$

from Lemma 9.9. Once we define

$$(x : A, w : \mathsf{Sub}\,B)\ \rho =_{\mathrm{def}} (\exists p : \mathsf{Sub}\,1)\big(\sigma \wedge \mathsf{coll}(p, w, (\_, y)\phi)\big) : \mathsf{prop}\,,$$

we have

$$(\forall x \in a)(\exists w : \mathsf{Sub}\,B)\rho$$

by the Strong Collection rule and (9.5). We derive that there is $u : \mathsf{Sub}(\mathsf{Sub}\,B)$ such that

$$\mathsf{coll}(x \in a, w \in u, (x, w)\rho)\,, \tag{9.6}$$

again by the Strong Collection rule. We get

$$(\forall x \in a)(\exists w : \mathsf{Sub}\,B)(\exists p : \mathsf{Sub}\,1)\big(\sigma \wedge (\forall\_ \in p)(\exists y \in w)\phi\big)$$

by (9.6) and the definition of ρ. This gives us

$$(\forall x \in a)(\exists p : \mathsf{Sub}\,1)\big(\sigma \wedge (\exists w \in u)(\exists y \in w)\phi\big)\,,$$

and thus we obtain

$$(\forall x \in a)J\big((\exists w \in u)(\exists y \in w)\phi\big)\,. \tag{9.7}$$

By (9.6) and the definition of ρ, it follows that

$$(\forall w \in u)(\exists x \in a)(\forall y \in w)\phi\,.$$

and therefore we get

$$(\forall w \in u)(\forall y \in w)(\exists x \in a)\phi\,. \tag{9.8}$$

Using Lemma 8.4 define

$$v =_{\text{def}} \bigcup u : \mathsf{Sub}\, B\,,$$

and observe that

$$(\forall x \in a)J(\exists y \in v)\phi \wedge (\forall y \in v)(\exists x \in a)\phi$$

by (9.7), (9.8) and Lemma 8.4. Discharging $u : \mathsf{Sub}(\mathsf{Sub}\, B)$ we get the desired conclusion. □

Theorem 9.11. *The j-interpretation of the Strong Collection rule is sound.*

Proof. The claim is a consequence of Lemma 9.10. □

To prove that the j-interpretation of the Subset Collection rule is sound we will introduce an additional hypothesis, that is intuitively related to the notion of set-presented nucleus given in Section 4.3.

Definition 9.12. We say that a topology j on P is **set-presented** if there exists $r : \mathsf{Sub}\,\mathsf{P}$ such that the judgement

$$(\forall p : \mathsf{P})\,\tau(jp) \equiv (\exists q \in r) q \leq p$$

is derivable. ◊

From now on we assume the Subset Collection rule. The next proposition is a type-theoretic version of Proposition 2.6.

Proposition 9.13. *Let A, B, C : type, let $(x : A, y : B, z : C)\ \phi$: prop and let $a : \mathsf{Sub}(\mathsf{Sub}\, A)$ and $b : \mathsf{Sub}\, B$. Then there exists $u : \mathsf{Sub}(\mathsf{Sub}\, B)$ such that the judgement*

$$(\forall w \in a)(\forall z : C)\big((\forall x \in w)(\exists y \in b)\psi \supset (\exists v \in u)\mathsf{coll}(w, v, (x,y)\psi)\big)$$

is derivable.

Proof. To prove the desired claim, it suffices to apply the Subset Collection rule and Lemma 8.4. ☐

The next proposition will be used in the proof of Lemma 9.15.

Lemma 9.14. *If j is a set-presented topology, then there exists $s : \mathsf{Sub}(\mathsf{Sub}\,1)$ such that, for $\phi : \mathsf{prop}$ the judgement*

$$J\phi \equiv (\exists p \in s)(\forall_{\_} \in p)\phi$$

is derivable.

Proof. The claim is a consequence of Lemma 8.6 and Lemma 8.7. ☐

Let us assume that the topology j is set-presented and that $s : \mathsf{Sub}(\mathsf{Sub}\,1)$ satisfies the property of Lemma 9.14.

Lemma 9.15. *Let $A, B, C : \mathsf{type}$, let $(x : A, y : B, z : C)\ \phi : \mathsf{prop}$, let $a : \mathsf{Sub}(A)$ and $b : \mathsf{Sub}(B)$. Then the judgement*

$$(\exists u'' : \mathsf{Sub}(\mathsf{Sub}\,B))(\forall z : C)\big((\forall x \in a)J(\exists y \in b)\phi \supset \\ (\exists v \in u'')\big((\forall x \in a)J(\exists y \in v)\phi\big) \wedge \big((\forall y \in v)(\exists x \in a)\phi\big)\big)$$

is derivable.

Proof. Let $a : \mathsf{Sub}(A)$ and $b : \mathsf{Sub}(B)$. Then, from Proposition 9.13 applied to $b : \mathsf{Sub}(B)$ and $s : \mathsf{Sub}(\mathsf{Sub}\,1)$ we get that there is $u : \mathsf{Sub}(\mathsf{Sub}\,B)$ such that

$$(\forall p \in s)(\forall x : A)(\forall z : C)\big((\forall_{\_} \in p)(\exists y \in b)\phi \ \supset\ (\exists v \in u)\mathsf{coll}(p, v, (\_, y)\phi)\big) \tag{9.9}$$

Define

$$(x : A, v : \mathsf{Sub}(\mathsf{Sub},\ B), z : C)\ \psi =_{\mathrm{def}} (\exists p \in s)\,\mathsf{coll}(p, v, (\_, y)\phi)$$

We get that there exists $u' : \mathsf{Sub}(\mathsf{Sub}(\mathsf{Sub}\,B))$ such that

$$(\forall z : C)\big((\forall x \in a)(\exists v \in u)\psi \supset (\exists w \in u')\,\mathsf{coll}(a, w, (x, v)\psi)\big) \tag{9.10}$$

by the Subset Collection rule. Let $z : C$ and assume that

$$(\forall x \in a)\ J(\exists y \in b)\phi$$

then we have

$$(\forall x \in a)(\exists p \in s)(\forall_{\_} \in p)(\exists y \in b)\phi$$

by Lemma 9.14 and therefore we obtain

$$(\forall x \in a)(\exists p \in s)(\exists v \in u)\mathsf{coll}(p, v, (\_, y)\phi)$$

by the implication in (9.9). By the definition of ψ we get

$$(\forall x \in a)(\exists v \in u)\psi\,.$$

It follows that

$$(\exists w \in u')(\forall x \in a)(\exists v \in w)\psi \wedge (\forall v \in w)(\exists x \in a)\psi$$

by the implication in (9.10). Let $w : \mathsf{Sub}(\mathsf{Sub}(B))$ and assume that

$$(\forall x \in a)(\exists v \in w)\psi \wedge (\forall v \in w)(\exists x \in a)\psi\,. \tag{9.11}$$

We have

$$(\forall x \in a)(\exists v \in w)(\exists p \in s)(\forall \_ \in p)(\exists y \in v)\phi$$

by the first conjunct in (9.11), and the definition of ψ. We therefore obtain

$$(\forall x \in a)(\exists p \in s)\big((\forall \_ \in p)(\exists y \in v)\phi\big) \tag{9.12}$$

Now observe that

$$(\forall v \in w)(\forall y \in v)(\exists x \in a)\phi \tag{9.13}$$

by the second conjunct of (9.11), and the definition of ψ. Now define

$$v =_{\text{def}} \bigcup w : \mathsf{Sub}\, B\,.$$

and observe that, discharging $w : \mathsf{Sub}(\mathsf{Sub}\, B)$, we get

$$(\exists v \in \{\bigcup w \mid w \in u'\})\ (\forall x \in a)\ J(\exists y \in v)\phi) \wedge (\forall y \in v)(\exists x \in a)\phi$$

by (9.12) and (9.13). We obtain the desired conclusion, with

$$u'' =_{\text{def}} \{\bigcup w \mid w \in u'\}\,,$$

as a consequence of Lemma 8.4 and Lemma 8.6. □

Theorem 9.16. *The j-interpretation of the Subset Collection rule is sound.*

Proof. Consequence of Lemma 9.15. □

We summarize the results obtained in this section in the next corollary, that is proved within $\mathbf{ML}(\mathbf{CZF}^-)$.

Corollary 9.17. *Let j be a topology.*

(i) *The j-interpretation of each rule of* $\mathbf{ML}(\mathbf{CZF}^-)$ *is sound.*

(ii) *Assuming the Subset Collection rule of* $\mathbf{ML}(\mathbf{CZF})$*, if j is set-presented, then the j-interpretation of the Subset Collection rule is sound.*

Proof. Part (i) follows by Proposition 9.8 and Theorem 9.11. Part (ii) follows by Theorem 9.16. □

9.5 Double-negation interpretation

As an application of the results just described we present a type-theoretic version of the double-negation interpretation. We define the double-negation topology as follows:

$$(x : \mathsf{P})\ jx =_{\text{def}} \dot{\neg}\,\dot{\neg}\ x : \mathsf{P}$$

where $\dot{\neg}\,x =_{\text{def}} x \mathbin{\dot{\supset}} \dot{\bot} : \mathsf{P}$, for $x : \mathsf{P}$. It is easy to prove that j is a topology. We refer to the j-interpretation determined by j as the double-negation interpretation.

Remark. Let us point out that the operator J determined by the double-negation topology seems to be logically equivalent to double negation only for small propositions, but not for arbitrary ones. In fact, for $\phi : \mathsf{prop}$ it holds

$$J\phi \equiv (\exists p : \mathsf{P})\big(\neg\neg\tau(p) \wedge \tau\, p \supset \phi\big)$$

where $\neg\phi =_{\text{def}} \phi \supset \bot$, for $\phi : \mathsf{prop}$. In general it will hold only that $J\phi$ implies $\neg\neg\phi$ but not vice versa.

It is natural to consider the following type-theoretic principle of restricted excluded middle, (**REM**), as in the following judgement:

$$(x : \mathsf{P})\ \tau(x)\ \vee\ \neg\tau(x)\,.$$

Theorem 9.18. *The double-negation interpretation of* $\mathbf{ML}(\mathbf{CZF}^-)$+**REM** *in* $\mathbf{ML}(\mathbf{CZF}^-)$ *is sound.*

Proof. Consequence of part (i) of Corollary 9.17. □

We can now consider a type-theoretic principle asserting that the double-negation topology is set presentable (**DNSP**):

$$(\exists r : \mathsf{Sub}\ \mathsf{P})(\forall x : \mathsf{P})(\,\neg\neg\tau(x) \equiv (\exists y \in r)\, y \leq x)\,.$$

Theorem 9.19. *The double-negation interpretation of* **ML**(**CZF**)+**REM** *in* **ML**(**CZF**)+ **DNSP** *is sound.*

Proof. Consequence of part (ii) of Corollary 9.17. □

Observe that the type theory **ML**(**CZF**) + **REM** and the set theory **CZF** + **REM** are mutually interpretable. Recall from Section 2.2 that the set theory **CZF**+**REM** has at least the proof-theoretic strength of second-order arithmetic and therefore the extension of the type theory **ML**(**CZF**) with **DNSP** pushes the proof-theoretic strength of the type theory above that of second-order arithmetic, and hence is fully impredicative.

Chapter 10

Conclusions

We have investigated generalised predicative intuitionistic formal systems both in set theory and in type theory. In set theory, we developed sheaf interpretations for CST and obtained first proof-theoretic applications thereof. In type theory, we introduced logic-enriched type theories and showed how they allow us to obtain a generalised type-theoretic interpretation of CST and to define reinterpretations of logic. We have also developed a significant fragment of formal topology in CST, so as to provide examples of sheaf interpretations.

The assumption of collection principles and the primitive treatment of logic deserve to be highlighted as two aspects that played the most prominent roles in our study both in set theory and in type theory. We leave to future research to investigate whether these aspects are relevant in other contexts as well. We expect, however, that a primitive treatment of logic in DTT may become very fruitful once logic-enriched type theories have been appropriately developed. This development may lead to applications in the computer-assisted formalisation of mathematics. We regard the treatment of non-extensional and extensional equalities in logic-enriched type theories as one of the most pressing issues to investigate.

Regarding CST, our development of sheaf interpretations opens the possibility of obtaining proof-theoretic results that were not within reach when we undertook the research described here. Regrettably, only a few examples of such results been presented. Further results should however be obtainable by direct application of the interpretations we developed, as we already discussed in Section 5.7. Sheaf interpretations seem also amenable of a more abstract development, using the ideas of AST. This would require as a preliminary step the isolation of axioms for small maps that correspond exactly to the axioms of CST. The only axiom, however, that is not covered by the existing development of AST seems to be Subset Collection.

Overall, we regard the research presented in this thesis as part of a long-term project, whose aims have been described in Section 1.4. An enormous amount of work remains to be done to complete the project, but we have taken some significant steps in that direction.

Appendix A

Axioms for set theories

Here we present all the axioms considered in CST, apart from the Regular Extension axiom (REA) for which we invite the reader to refer to [7, Section 5.2].

A.1 Structural axioms

Extensionality. For all sets a and b

$$(\forall x)(x \in a \leftrightarrow x \in b) \rightarrow a = b\,.$$

Set induction. For arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\forall x)((\forall y \in x)\,\phi[y/x] \rightarrow \phi) \rightarrow (\forall x)\phi\,.$$

Foundation. For arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\exists x)\phi \rightarrow (\exists x)\big(\phi \wedge (\forall y \in x)\neg\phi[y/x]\big)\,.$$

A.2 Set existence axioms

Pairing. For all sets a and b

$$(\exists u)(\forall x)\big(x \in u \leftrightarrow x = a \vee x = b\big)\,.$$

Union. For all sets a

$$(\exists u)(\forall x)\big(x \in u \leftrightarrow (\exists y \in a)x \in y\big)\,.$$

Infinity.

$$(\exists u)\big((\exists x)x \in u \wedge (\forall x \in u)(\exists y \in u)x \in y\big)\,.$$

Power set. For all sets a

$$(\exists u)(\forall x)\big(x \in u \leftrightarrow (\forall y \in x)y \in a\big)\,.$$

A.3 Separation axioms

Restricted separation. For all sets a and for restricted formulas θ of $\mathcal{L}^{(V)}$

$$(\exists u)(\forall x)\big(x \in u \leftrightarrow x \in a \wedge \theta\big)\,.$$

Full separation. For all sets a and for arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\exists u)(\forall x)\big(x \in u \leftrightarrow x \in a \wedge \phi\big)\,.$$

A.4 Collection axioms

For sets a, u and for an arbitrary formula ϕ of $\mathcal{L}^{(V)}$, define

$$\mathrm{coll}(x \in a, y \in u, \phi) =_{\mathrm{def}} (\forall x \in a)(\exists y \in u)\phi \wedge (\forall y \in u)(\exists x \in a)\phi\,.$$

Replacement. For all sets a and for arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\forall x \in a)(\exists! y)\phi(x, y) \rightarrow (\exists u)\mathrm{coll}(x \in a, y \in u, \phi)\,.$$

Strong collection. For all sets a and for arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\forall x \in a)(\exists y)\phi(x, y) \rightarrow (\exists u)\mathrm{coll}(x \in a, y \in u, \phi)\,.$$

Subset collection. For all sets a and b, and for arbitrary formulas ϕ of $\mathcal{L}^{(V)}$

$$(\exists v)(\forall z)\big((\forall x \in a)(\exists y \in b)\phi \rightarrow (\exists u \in v)\mathrm{coll}(x \in a, y \in u, \phi)\big)\,.$$

Appendix B

Rules for dependent type theories

B.1 General rules

Assumption rule.

$$\frac{(\Gamma, \Delta)\ J \qquad A : \mathsf{type}}{(\Gamma, x : A, \Delta)\ x : A}\ x \notin \mathrm{FV}(\Gamma) \cup \mathrm{FV}(\Delta)$$

From now on we suppress mention of a context that is common to both the premisses and the conclusion of a rule.

Equality rules.

$$\frac{A : \mathsf{type}}{A = A : \mathsf{type}} \qquad \frac{A_1 = A_2 : \mathsf{type}}{A_2 = A_1 : \mathsf{type}} \qquad \frac{A_1 = A_2 : \mathsf{type} \qquad A_2 = A_3 : \mathsf{type}}{A_1 = A_3 : \mathsf{type}}$$

$$\frac{a : A}{a = a : A} \qquad \frac{a_1 = a_2 : A}{a_2 = a_1 : A} \qquad \frac{a_1 = a_2 : A \qquad a_2 = a_3 : A}{a_1 = a_3 : A}$$

$$\frac{a : A_1 \qquad A_1 = A_2}{a : A_2} \qquad \frac{a_1 = a_2 : A_1 \qquad A_1 = A_2}{a_1 = a_2 : A_2}$$

Substitution rule.

$$\frac{(x : A, \Delta)\ \mathcal{B} \qquad a : A}{(\Delta[a/x])\ \mathcal{B}[a/x]}$$

Congruence rules.

$$\frac{(x : A, \Delta)\ C : \mathsf{type} \qquad a_1 = a_2 : A}{(\Delta[a_1/x])\ C[a_1/x] = C[a_2/x] : \mathsf{type}} \qquad \frac{(x : A, \Delta)\ c : C \qquad a_1 = a_2 : A}{(\Delta[a_1/x])\ c[a_1/x] = c[a_2/x] : C[a_1/x]}$$

Special congruence rules. For ∇ that is either Π, Σ or W:

$$\frac{(x : A)\ B_1 = B_2 : \mathsf{type}}{(\nabla x : A)B_1 = (\nabla x : A)B_2 : \mathsf{type}} \qquad \frac{(x : A)\ b_1 = b_2 : B}{(\lambda x : A)b_1 = (\lambda x : A)b_2 : (\Pi x : A)B}$$

Analogous rules should also be formulated for other symbols such as $\mathsf{r_N}$, split and $\mathsf{r_W}$, but we omit for brevity.

B.2 Type rules

B.2.1 Non-dependent types

0-rules.

$$\mathsf{0} : \mathsf{type} \tag{0-form.}$$

$$\frac{(z : \mathsf{0})\ C : \mathsf{type} \qquad e : \mathsf{0}}{\mathsf{r_0}(e) : C[e/z]} \tag{0-elim.}$$

1-rules.

$$\mathsf{1} : \mathsf{type} \tag{1-form.}$$

$$0_1 : \mathsf{1} \tag{1-intro.}$$

$$\frac{(z : \mathsf{1})\ C : \mathsf{type} \qquad c : C[0_1/z] \qquad e : \mathsf{1}}{\mathsf{r_1}(c, e) : C[e/z]} \tag{1-elim.}$$

$$\frac{(z : \mathsf{1})\ C : \mathsf{type} \qquad c : C[0_1/z]}{\mathsf{r_1}(c, 0_1) = c : C[0_1/z]} \tag{1-comp.}$$

2-rules.

$$\mathsf{2} : \mathsf{type} \tag{2-form.}$$

$$1_2 : \mathsf{2} \qquad\qquad 2_2 : \mathsf{2} \tag{2-Intr.}$$

Define the judgements

- $J =_{\text{def}} (z : 2)\ C : \mathsf{type}$
- $J_1 =_{\text{def}} c_1 : C[1_2/z]$
- $J_2 =_{\text{def}} c_2 : C[2_2/z]$

$$\frac{J \qquad J_1 \qquad J_2 \qquad e : 2}{\mathsf{r}_2(c_1, c_2, e) : C[e/z]} \tag{2-elim.}$$

$$\frac{J \qquad J_1 \qquad J_2}{\mathsf{r}_2(c_1, c_2, 1_2) = c_1 : C[1_2/z]} \qquad \frac{J \qquad J_1 \qquad J_2}{\mathsf{r}_2(c_1, c_2, 2_2) = c_2 : C[2_2/z]} \tag{2-comp.}$$

N-rules.

$$\mathsf{N} : \mathsf{type} \tag{N-form.}$$

$$0 : \mathsf{N} \qquad \frac{a : \mathsf{N}}{\mathsf{succ}(a) : \mathsf{N}} \tag{N-intro.}$$

Define the judgements:

- $J =_{\text{def}} (z : \mathsf{N})\ C : \mathsf{type}$,
- $J_0 =_{\text{def}} c : C[0/z]$,
- $J_{\mathsf{succ}} =_{\text{def}} (x : \mathsf{N}, y : C[x/z])\ d : C[\mathsf{succ}(x)/z]$.

and the expression:

- $H(e) =_{\text{def}} \mathsf{r}_{\mathsf{N}}(c, (x, y)d, e)$.

$$\frac{J \qquad J_0 \qquad J_{\mathsf{succ}} \qquad e : \mathsf{N}}{H(e) : C[e/z]} \tag{N-elim.}$$

$$\frac{J \qquad J_0 \qquad J_{\mathsf{succ}}}{H(0) = c : C[0/z]} \qquad \frac{J \qquad J_0 \qquad J_{\mathsf{succ}} \qquad a : \mathsf{N}}{H(\mathsf{succ}(a)) = d[a, H(a)/x, y] : C[\mathsf{succ}(a)/z]} \tag{N-comp.}$$

B.2.2 Dependent types

R_2-rules.

$$\frac{A_1 : \mathsf{type} \qquad A_2 : \mathsf{type} \qquad e : \mathsf{2}}{\mathsf{R}_2(A_1, A_2, e) : \mathsf{type}} \qquad (\mathsf{R}_2\text{-form.})$$

$$\mathsf{R}_2(A_1, A_2, 1_2) = A_1 \qquad \mathsf{R}_2(A_1, A_2, 2_2) = A_2 \qquad (\mathsf{R}_2\text{-Conv.})$$

Σ-rules.

$$\frac{A : \mathsf{type} \qquad (x : A)\ B : \mathsf{type}}{(\Sigma x : A)B : \mathsf{type}} \qquad (\Sigma\text{-form.})$$

$$\frac{a : A \qquad b : B[a/x]}{\mathsf{pair}(a, b) : (\Sigma x : A)B} \qquad (\Sigma\text{-intro.})$$

Define the judgements

- $J =_{\text{def}} (z : (\Sigma x : A)B)\ C : \mathsf{type}$,
- $J_{\mathsf{pair}} =_{\text{def}} (x : A, y : B)\ c : C[\mathsf{pair}(x, y)/z]$,

and the expression

- $H(e) =_{\text{def}} \mathsf{split}((x, y)c, e)$.

$$\frac{J \qquad J_{\mathsf{pair}} \qquad e : (\Sigma x : A)B}{H(e) : C[e/z]} \qquad (\Sigma\text{-elim.})$$

$$\frac{J \qquad J_{\mathsf{pair}} \qquad a : A \qquad b : B[a/x]}{H(\mathsf{pair}(a, b)) = c[a, b/x, y] : C[\mathsf{pair}(a, b)/z]} \qquad (\Sigma\text{-comp.})$$

Π-rules.

$$\frac{A : \mathsf{type} \qquad (x : A)\ B : \mathsf{type}}{(\Pi x : A)B : \mathsf{type}} \qquad (\Pi\text{-form.})$$

$$\frac{(x : A)\ b : B}{(\lambda x : A)b : (\Pi x : A)B} \qquad (\Pi\text{-intro.})$$

$$\frac{f : E \qquad a : A}{\mathsf{app}(f, a) : B[a/x]} \qquad (\Pi\text{-elim.})$$

$$\frac{(x : A)\ b : B \qquad a : A}{\mathsf{app}((\lambda x : A)b, a) = b[a/x] : B[a/x]} \qquad (\Pi\text{-comp.})$$

W-rules.

$$\frac{A : \mathsf{type} \qquad (x : A)\ B : \mathsf{type}}{(Wx : A)B : \mathsf{type}} \tag{W-form.}$$

$$\frac{a : A \qquad b : B[a/x] \to (Wx : A)B}{\mathsf{sup}(a, b) : (Wx : A)B} \tag{W-intro.}$$

Define the expression

- $E =_{\text{def}} (Wx : A)B$,

the judgements

- $J =_{\text{def}} (z : E)\ C : \mathsf{type}$,
- $J_{\mathsf{sup}} =_{\text{def}} (x : A, u : B \to E, v : (\Pi y : B)C[\mathsf{app}(u, y)/z])\ c : C[\mathsf{sup}(x, u)/z]$

and the expression

- $H(e) =_{\text{def}} \mathsf{r_W}((x, u, v)c, e)$.

$$\frac{J \qquad J_{\mathsf{sup}} \qquad e : E}{H(e) : C[e/z]} \tag{W-elim.}$$

$$\frac{J \qquad J_{\mathsf{sup}} \qquad a : A \qquad b : B[a/x] \to E}{H(e) = c[a, b, (\lambda y : B[a/x])H(\mathsf{app}(b, y))/x, u, v] : C[\mathsf{sup}(a, b)/z]} \tag{W-comp.}$$

B.2.3 Type universe

$$\mathsf{U} : \mathsf{type} \tag{U-form.}$$

$$\frac{a : \mathsf{U}}{\mathsf{T}\,a : \mathsf{type}} \tag{U-elim.}$$

In the next rules:

- q is one of the following: $\mathsf{0}$, $\mathsf{1}$, $\mathsf{2}$, N.
- ∇ is either Σ or Π.

$$\dot{q} : \mathsf{U} \qquad\qquad \frac{a : \mathsf{U} \qquad (x : \mathsf{T}\,a)\ b : \mathsf{U}}{(\dot{\nabla} x : a)b : \mathsf{U}} \tag{U-intro.}$$

$$\mathsf{T}\,\dot{q} = q : \mathsf{type} \qquad \frac{a : \mathsf{U} \qquad (x : \mathsf{T}\,a)\ b : \mathsf{U}}{\mathsf{T}(\dot{\nabla} x : a)b = (\nabla x : \mathsf{T}\,a)\,\mathsf{T}\,b : \mathsf{type}} \tag{U-comp.}$$

B.2.4 Proposition universe

$$\mathsf{P} : \mathsf{type} \qquad \text{(P-form.)}$$

$$\frac{p : \mathsf{P}}{\tau\, p : \mathsf{prop}} \qquad \text{(P-elim.)}$$

In the rules that follow:

- q is either $\top$ or $\bot$,
- $\star$ is one of the following: $\wedge, \vee, \supset$,
- ∇ is either $\forall$ or $\exists$.

$$\dot{q} : \mathsf{P} \qquad \frac{p_1 : \mathsf{P} \qquad p_2 : \mathsf{P}}{p_1 \,\dot{\star}\, p_2 : \mathsf{P}} \qquad \frac{a : \mathsf{U} \qquad (x : \mathsf{T}\, a)\ p : \mathsf{P}}{(\dot{\nabla} x : a)p : \mathsf{P}} \qquad \text{(P-intro.)}$$

$$\tau\, \dot{q} \equiv q \qquad \frac{p_1 : \mathsf{P} \qquad p_2 : \mathsf{P}}{\tau(p_1 \,\dot{\star}\, p_2) \equiv \tau\, p_1 \star \tau\, p_2} \qquad \frac{a : \mathsf{U} \qquad (x : \mathsf{T}\, a)\ p : \mathsf{P}}{\tau(\dot{\nabla} x : a)p \equiv (\nabla x : \mathsf{T}\, a)\, \tau\, p} \qquad \text{(P-comp.)}$$

B.3 Rules for ML(CZF)

Π^--rules

$$\frac{a : \mathsf{U} \qquad (x : \mathsf{T}\, a)\ B : \mathsf{type}}{(\Pi^- x : a)B : \mathsf{type}} \qquad (\Pi^-\text{-formation})$$

The introduction, elimination and computation rule are identical to the ones for Π-types, as given on page 190.

W^--rules

$$\frac{A : \mathsf{type} \qquad (x : A)\ b : \mathsf{U}}{(W^- x : A)b : \mathsf{type}} \qquad (W^-\text{-formation})$$

The introduction rules are analogous to the ones for W-types on page 191. Let $A : \mathsf{type}$ and $(x : A)\ b : \mathsf{U}$. Define the expressions:

- $B =_{\mathrm{def}} \mathsf{T}\, b$,
- $E =_{\mathrm{def}} (W^- x : A)b$.

Let $(z_1, z_2 : E)\ C : \mathsf{type}$ and define the contexts:

- $\Gamma_1 =_{\text{def}} x_1 : A, u_1 : B[x_1/x] \to E\,,$
- $\Gamma_2 =_{\text{def}} x_2 : A, u_2 : B[x_2/x] \to E\,,$
- $\Delta =_{\text{def}} v : (\Pi y_1 : B[x_1/x])(\Pi y_2 : B[x_2/x])C[\mathsf{app}(u_1, y_1), \mathsf{app}(u_2, y_2)/z_1, z_2]\,.$

Define the judgements:

- $J =_{\text{def}} (z_1, z_2 : E)\; C : \mathsf{type}$
- $J_{\mathsf{sup}} =_{\text{def}} (\Gamma_1, \Gamma_2, \Delta)\; d : C[\mathsf{sup}[(x_1, u_1), \mathsf{sup}(x_2, u_2)/z_1, z_2]$

and the expression

- $H(e_1, e_2) =_{\text{def}} \mathsf{r}_{\mathsf{2W}}((x_1, u_1, x_2, u_2, v)d, e_1, e_2)$

$$\frac{J \qquad J_{\mathsf{sup}} \qquad e_1 : E \qquad e_2 : E}{H(e_1, e_2) : C[e_1, e_2/z_1, z_2]} \qquad \text{(Double } W^{-}\text{-elim.)}$$

The double W^{-}-elimination rule is analogous to the one for W-types.

Appendix C

Axioms for small maps

Let $\mathcal{E}$ be a regular category with finite coproducts that are disjoint and stable. The following axioms refer to objects and maps of $\mathcal{E}$.

C.1 Basic axioms

(A1) Any isomorphism is small, small maps are closed under composition.

(A2) In a pullback diagram

$$\begin{array}{ccc} B & \longrightarrow & A \\ {\scriptstyle G}\downarrow & & \downarrow{\scriptstyle F} \\ Y & \longrightarrow & X \end{array}$$

if F is small then G is small.

(A3) In a pullback

$$\begin{array}{ccc} B & \longrightarrow & A \\ {\scriptstyle G}\downarrow & & \downarrow{\scriptstyle F} \\ Y & \underset{P}{\twoheadrightarrow} & X \end{array}$$

if P is epi and G is small then F is small.

(A4) The maps $0 \longrightarrow 1$ and $1 + 1 \longrightarrow 1$ are small.

(A5) If $A \longrightarrow X$ and $B \longrightarrow Y$ are small, then so is $A + B \longrightarrow X + Y$

(A6) In a commuting diagram

$$\begin{array}{ccc} B & \xrightarrow{P} & A \\ & {\scriptstyle G}\searrow \quad \swarrow{\scriptstyle F} & \\ & X & \end{array}$$

If P is epi and G is small, then F is small.

C.2 Power classes

In the next axiom $\mathcal{P}$ and $\ni$ are definable operations on objects of $\mathcal{E}$ such that if A is an object of $\mathcal{E}$, then $\mathcal{P}(A)$ is also an object of $\mathcal{E}$ and

$$\ni_A \rightarrowtail \mathcal{P}(A) \times A$$

is a $\mathcal{P}(A)$-indexed family of small subobjects of A in the sense of Definition 2.1.

(P1) Let A be an object of $\mathcal{E}$. For all objects I in $\mathcal{E}$ and all I-indexed families of small subobjects of A $R \rightarrowtail I \times A$ there exists a unique map $I \xrightarrow{F} \mathcal{P}(A)$ such that the diagram

$$\begin{array}{ccc} R & \longrightarrow & \ni_A \\ \downarrow & & \downarrow \\ I \times A & \xrightarrow[F \times \mathrm{Id}_A]{} & \mathcal{P}(A) \times A \end{array}$$

is a pullback.

Bibliography

[1] P. Aczel. Extending the topological interpretation to Constructive Set Theory. Manuscript notes, 1977.

[2] P. Aczel. The type theoretic interpretation of Constructive Set Theory. In A. MacIntyre, L. Pacholski, and J. Paris, editors, *Logic Colloquium '77*, pages 55 – 66. North-Holland, 1978.

[3] P. Aczel. The type theoretic interpretation of Constructive Set Theory: choice principles. In Troelstra and van Dalen [83], pages 1 – 40.

[4] P. Aczel. The type theoretic interpretation of Constructive Set Theory: inductive definitions. In R. Barcan Marcus, G.J.W. Dorn, and P. Weinegartner, editors, *Logic, Methodology and Philosophy of Science VII*, pages 17 – 49. North-Holland, 1986.

[5] P. Aczel. On relating type theories and set theories. In T. Altenkirch, W. Naraschewski, and B. Reus, editors, *Types for Proofs and Programs*, volume 1257 of *Lecture Notes in Computer Science*. Springer, 2000.

[6] P. Aczel and N. Gambino. Collection principles in Dependent Type Theory. In P. Callaghan, Z. Luo, J. McKinna, and R. Pollack, editors, *Types for Proofs and Programs*, volume 2277 of *Lecture Notes in Computer Science*, pages 1 – 23. Springer, 2002.

[7] P. Aczel and M. Rathjen. Notes on Constructive Set Theory. Technical Report 40, Mittag-Leffler Institute, The Swedish Royal Academy of Sciences, 2001. Available from the first author's web page at `http://www.cs.man.ac.uk/~petera/papers.html`.

[8] S. Awodey, C. Butz, T. Streicher, and A. K. Simpson. Relating toposes and first-order set theories. Draft paper, 2002.

[9] G. Battilotti and G. Sambin. A uniform presentation of sup-lattices, quantales and

frames by means of infinitary preordered sets, pretopologies and formal topologies. Technical Report 19, Department of Mathematics, University of Padua, 1993.

[10] M. J. Beeson. *Foundations of Constructive Mathematics*. Springer, 1985.

[11] J. L. Bell. *Boolean-valued Models and Independence Proofs in Set Theory*. Clarendon Press, 1977.

[12] A. Boileau and A. Joyal. La logique des topos. *Journal of Symbolic Logic*, 46(1):6 – 16, 1981.

[13] D. Bridges and F. Richman. *Varieties of Constructive Mathematics*. London Mathematical Society Lecture Notes. Cambridge University Press, 1987.

[14] A. Carboni. Some free constructions in realizability and proof theory. *Journal of Pure and Applied Algebra*, 103:117 – 148, 1995.

[15] A. Carboni and G. Rosolini. Locally Cartesian closed exact completions. *Journal of Pure and Applied Algebra*, 154:103 – 116, 2000.

[16] A. Carboni and E. M. Vitale. Regular and exact completions. *Journal of Pure and Applied Algebra*, 125:79 – 116, 1998.

[17] T. Coquand. Notions invariant by change of base. Available from the author's web page at: `http://www.cs.chalmers.se/~coquand`.

[18] T. Coquand. Metamathematical investigations of a Calculus of Constructions. In P. Odifreddi, editor, *Logic and Computer Science*, pages 91 –122. Academic Press, 1990.

[19] T. Coquand and E. Palmgren. Intuitionistic choice and classical logic. *Archive for Mathematical Logic*, 39(1):53 – 74, 2000.

[20] T. Coquand, G. Sambin, J. M. Smith, and S. Valentini. Inductively generated formal topologies. Submitted for publication. Available from the first author's web page at: `http://www.cs.chalmers.se/~coquand`.

[21] G. Curi. Compact Hausdorff spaces are data types. Draft paper, 2002.

[22] G. Curi. The points of (locally) compact regular formal topologies. In Peter Schuster, Ulrich Berger, and Horst Osswald, editors, *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*, pages 39 – 54. Kluwer, 2002.

[23] A. G. Dragalin. *Mathematical Intuitionism — Introduction to Proof Theory*, volume 67 of *Translations of Mathematical Monographs.* American Mathematical Society, 1980.

[24] M. Dummett. *Elements of Intuitionism.* Oxford University Press, second edition, 2000.

[25] M. H. Escardó and A. K. Simpson. A universal characterization of the closed Euclidean interval (extended abstract). In *16th Annual IEEE Symposium on Logic in Computer Science*, pages 115 – 125. IEEE Press, 2001.

[26] S. Feferman. Formal theories for transfinite iterations of generalized inductive definitions and some subsystems of analysis. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Intuitionism and Proof Theory*, pages 303 – 325. North-Holland, 1970.

[27] S. Feferman. Hilbert's program relativized: proof theoretical and foundational reductions. *Journal of Symbolic Logic*, 53(2):364 – 384, 1988.

[28] S. Feferman. *In the Light of Logic.* Oxford University Press, 1998.

[29] M. P. Fourman. Sheaf models for set theory. *Journal of Pure and Applied Algebra*, 19:91 – 101, 1980.

[30] M. P. Fourman and R. J. Grayson. Formal spaces. In Troelstra and van Dalen [83], pages 107 – 122.

[31] M. P. Fourman and J. M. E. Hyland. Sheaf models for analysis. In Fourman et al. [32].

[32] M. P. Fourman, C. J. Mulvey, and D. S. Scott, editors. *Applications of Sheaves*, volume 753 of *Lecture Notes in Mathematics.* Springer, 1979.

[33] M. P. Fourman and A. Scedrov. The world's weakeast axiom of choice fails. *Manuscripta Mathematica*, 38:325 – 332, 1982.

[34] H. M. Friedman. The consistency of classical set theory relative to a set theory with intuitionistic logic. *Journal of Symbolic Logic*, 38:315–319, 1973.

[35] H. M. Friedman. Set theoretic foundations of constructive analysis. *Annals of Mathematics*, 105:1–28, 1977.

[36] N. Gambino. Types and sets — A study on the jump to full impredicativity, 1999. Laurea dissertation, Department of Pure and Applied Mathematics, University of Padua.

[37] N. Gambino and P. Aczel. Frame-valued semantics for Constructive Set Theory. Technical Report 39, Institut Mittag-Leffler, The Swedish Royal Academy of Sciences, 2001.

[38] R. J. Grayson. Heyting-valued models for Intuitionistic Set Theory. In Fourman et al. [32], pages 402 – 414.

[39] R. J. Grayson. Forcing in intuitionistic systems without power-set. *Journal of Symbolic Logic*, 48(3):670–682, 1983.

[40] E. R. Griffor and M. Rathjen. The strength of some Martin-Löf type theories. *Archive for Mathematical Logic*, 33:347 – 385, 1994.

[41] M. Hofmann. On the interpretation of type theory in locally Cartesian closed categories. In L. Pacholski and J. Tiuryn, editors, *Computer Science Logic 1994*, volume 933 of *Lecture Notes in Computer Science*, pages 427 – 441. Springer, 1995.

[42] M. Hofmann. *Extensional Constructs in Intensional Type Theory*. Springer, 1997.

[43] M. Hofmann. Syntax and semantics of dependent types. In A. M. Pitts and P. Dybjer, editors, *Semantics of Logics of Computation*, Publications of the Newton Institute, pages 79 – 130. Cambridge University Press, 1997.

[44] W. A. Howard. The formulae-as-types notion of construction. In J. P. Hindley and J. R. Seldin, editors, *To H. B. Curry: essays on combinatory logic, lambda calculus, and formalism*, pages 480 – 490. Academic Press, 1980.

[45] J. M. E. Hyland and A. M. Pitts. The theory of constructions: categorical semantics and topos-theoretic models. In J. W. Gray and A. Scedrov, editors, *Categories in Computer Science and Logic*, Contemporary Mathematics, pages 137 – 199. American Mathematical Society, 1989.

[46] P. T. Johnstone. *Stone Spaces*. Cambridge University Press, 1982.

[47] P. T. Johnstone. The point of pointless topology. *Bullettin of the American Mathematical Society*, 8:41 – 53, 1983.

[48] P. T. Johnstone. The art of pointless thinking: a student's guide to the category of locales. In H. Herrlich and H.-E. Porst, editors, *Category Theory at Work*, pages 85 – 107. Heldermann Verlag, 1991.

[49] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*. Oxford University Press, 2002.

[50] A. Joyal and I. Moerdijk. A completeness theorem for open maps. *Annals of Pure and Applied Logic*, 70:51 – 86, 1994.

[51] A. Joyal and I. Moerdijk. *Algebraic Set Theory*, volume 220 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, 1995.

[52] S. C. Kleene. On the interpretation of intuitionistic number theory. *Journal of Symbolic Logic*, 10:109 – 124, 1945.

[53] K. Kunen. *Set Theory — An Introduction to Independence Proofs*. North-Holland, 1980.

[54] F. W. Lawvere. Quantifiers as sheaves. In *Proceedings International Congress on Mathematics*, pages 1506 – 1511. Gauthier-Villars, 1971.

[55] Z. Luo. *Computation and Reasoning — A Type Theory for Computer Science*. Oxford University Press, 1994.

[56] S. MacLane and I. Moerdijk. *Sheaves in Geometry and Logic — A First Introduction to Topos Theory*. Springer, 1992.

[57] M. E. Maietti. *The type theory of categorical universes*. PhD thesis, Department of pure and applied mathematics, Univesity of Padua, 1998.

[58] M. E. Maietti. Modular correspondence between dependent type theories and categorical universes. Technical Report 44, Institut-Mittag Leffler, The Royal Swedish Academy of Sciences, 2001.

[59] P. Martin-Löf. *Intuitionistic Type Theory — Notes by G. Sambin of a series of lectures given in Padua, June 1980*. Bibliopolis, 1984.

[60] M. Menni. *Exact completions and toposes*. PhD thesis, University of Edinburgh, 2000.

[61] I. Moerdijk and E. Palmgren. Wellfounded trees in categories. *Annals of Pure and Applied Logic*, 104:189 – 218, 2000.

[62] I. Moerdijk and E. Palmgren. Type theories, toposes and Constructive Set Theory: predicative aspects of AST. *Annals of Pure and Applied Logic*, 114(1-3):155–201, 2002.

[63] J.R. Myhill. Constructive Set Theory. *Journal of Symbolic Logic*, 40(3):347–382, 1975.

[64] B. Nordström, K. Petersson, and J. M. Smith. *Programming in Martin-Löf's Type Theory — An Introduction*. Oxford University Press, 1990.

[65] B. Nordström, K. Petersson, and J. M. Smith. Martin-Löf Type Theory. In S. Abramski, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 5. Oxford University Press, 2000.

[66] A. M. Pitts. Categorical Logic. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 5. Oxford University Press, 2000.

[67] M. Rathjen, E. R. Griffor, and E. Palmgren. Inaccessibility in Constructive Set Theory and Type Theory. *Annals of Pure and Applied Logic*, 94:181 – 200, 1998.

[68] G. Sambin. Intuitionistic formal spaces. In D. Skordev, editor, *Mathematical Logic and its Applications*, pages 87 – 204. Plenum, 1987.

[69] G. Sambin. Some points in formal topology. *Theoretical Computer Science*, To appear. Available from the author's web page at: `http://www.math.unipd.it/~sambin`.

[70] G. Sambin and S. Gebellato. A preview of the basis picture: a new perspective on formal topology. In T. Altenkirch, W. Naraschewski, and B. Reus, editors, *Types for Proofs and Programs*, volume 1257 of *Lecture Notes in Computer Science*, pages 194 – 207. Springer, 2000.

[71] G. Sambin and S. Valentini. Building up a toolbox for Martin-Löf's type theory: subset theory. In G. Sambin and J. M. Smith, editors, *Twenty-five Years of Constructive Type Theory*, pages 221–244. Oxford University Press, 1998.

[72] A. Scedrov. Consistency and independence results in Intuitionistic Set Theory. In F. Richman, editor, *Constructive Mathematics*, volume 873 of *Lecture Notes in Mathematics*, pages 54 – 86. Springer, 1981.

[73] A. Scedrov. Independence of the fan theorem in the presence of continuity principles. In Troelstra and van Dalen [83], pages 435 – 442.

[74] A. Scedrov. Intuitionistic Set Theory. In L.A. Harrington, M.D. Morley, A. Scedrov, and S.G. Simpson, editors, *Harvey Friedman's Research on the Foundations of Mathematics*, pages 257 – 284. North-Holland, 1985.

[75] D. S. Scott. Extending the topological interpretation to intuitionistic analysis. *Compositio Mathematica*, 20:194 – 210, 1968.

[76] D. S. Scott. Constructive validity. In M. Laudet and D. Lacombe, editors, *Symposium on Automated Demonstration*, volume 125 of *Lecture Notes in Mathematics*, pages 237 – 275. Springer, 1970.

[77] D. S. Scott. Extending the topological interpretation to intuitionistic analysis, II. In J. Myhill, A. Kino, and R. Vesley, editors, *Intuitionism and Proof Theory*, pages 235 – 255. North-Holland, 1970.

[78] D. S. Scott. Category-theoretic models for Intuitionistic Set Theory. Manuscript slides of a talk given at Carnagie-Mellon University, 1985.

[79] R.A.G. Seely. Locally Cartesian closed categories and type theory. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95:33 – 48, 1984.

[80] H. Simmons. A framework for topology. In A. MacIntyre, L. Pacholski, and J. Paris, editors, *Logic Colloqiuum '77*, pages 239 – 251. North-Holland, 1978.

[81] A. K. Simpson. Elementary axioms for the category of classes. In *14th Annual IEEE Symposium on Logic in Computer Science*, pages 77 – 85. IEEE Press, 1999.

[82] A. S. Troelstra. Realizability. In S. R. Buss, editor, *Handbook of Proof Theory*. North-Holland, 1998.

[83] A. S. Troelstra and D. van Dalen, editors. *The L. E. J. Brouwer Centenary Symposium*. North-Holland, 1982.

[84] J. K. Truss. *Foundations of Mathematical Analysis*. Oxford University Press, 1997.

[85] S. Valentini. On the formal points of the formal topology of the binary tree. *Archive for Mathematical Logic*, To appear. Available from the author's web page: `http://www.math.unipd.it/~silvio`.