Ludwig Maximilian University of Munich
Munich Center for Mathematical Philosophy

Constructivity in Homotopy Type Theory

Author:
Maximilian Doré

Supervisors:
Prof. Dr. Dr. Hannes Leitgeb
Prof. Steve Awodey, PhD

Munich, August 2019

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Logic and Philosophy of Science

Contents

Chapter 1

Introduction

In the seminal HoTT Book (2013), it is maintained that constructive reasoning is not necessary when formalizing mathematics in homotopy type theory (HoTT):

> It is worth emphasizing that univalent foundations does not *require* the use of constructive or intuitionistic logic.
>
> — HoTT Book (2013):p. 13

The law of excluded middle (LEM) and the axiom of choice can be assumed consistently in appropriate form and hence, HoTT is not committed to constructive mathematics. Since the backbone of HoTT is *intuitionistic* type theory developed by Per Martin-Löf, the question arises which repercussions the combination of an intuitionistic formal theory with LEM has for the philosophy of mathematics — the LEM has been the bone of contention for intuitionistic mathematicians from the beginning and led to the development of a distinctively anti-realist stance on the nature of mathematics, which seems in conflict with employing the LEM. This question provided the starting point for this thesis to investigate the different notions of constructivity used in HoTT and how classical reasoning relates to these notions.

It is uncontroversial among many mathematicians who find the law of excluded middle not problematic that nevertheless, a constructive proof is preferable to a proof not yielding the witness of an existential or disjunctive statement. This view is shared by some of the researchers working on constructive mathematics, this pragmatic position has been termed "liberal constructivism" in the literature by Billinge (2003). The roots of this liberal standpoint can be seen to go back to Errett Bishop, who revived constructive mathematics in Bishop (1967) by showing that constructive reasoning does not have to be cumbersome. Bishop maintained that using the LEM where necessary is not incoherent, but did not present a philosophy of mathematics underpinning his approach, as Billinge attests. The author is not aware of a systematic account that explains what liberal constructivism means for the semantics, epistemology and ontology of mathematics. In particular, such an account has to answer the following questions: Is the constructively proven part and the classically proven part of mathematics about the same subject matter? How is the status of constructively proven theorems different from theorems that are only classically valid?

When answering these questions, one has to lay down how the logical constants are understood. In the intuitionistic tradition, the meaning of an expression is determined by its use, whereas classically minded logicians specify the meaning of the logical constants by investigating the conditions under which they are true. A reconciliation of both views seems necessary when employing the law of excluded middle in an intuitionistic formal system.

In order to formulate such a reconciliation we need to understand what has changed since the initial development of intuitionistic logic by Heyting and Kolmogorov. We will see in the course of this thesis that intuitionistic type theory has distinctively more features than what is commonly called intuitionistic logic — the HoTT Book (2013) introduces for this purpose the distinction between constructivity in the "algorithmic" sense and constructivity in the "intuitionistic" sense. We will argue that algorithmic constructivity can be seen as standing in the tradition of the intuitionistic approach to logic, but also introduces new concepts that cannot be explained by the BHK meaning explanations of intuitionistic logic. In particular, modern constructive type theory treats proofs as algorithms and assigns them a central role in the system as first-order objects. Furthermore, intuitionistic type theory presents a highly intricate treatment of identity. This treatment of identity remains mysterious to this day, we want to highlight the features of it and appreciate why they come unexpected even for the creator of intuitionistic type theory, Per Martin-Löf.

Both intuitionists and realists are faced with a fundamental problem: Intuitionists need to explain how subjective mental constructions can be communicated intersubjectively and how mathematics is so successful if it is solely a product of the mind, whereas realists need to explain how it is that we can interact with "abstract" entities. This thesis can be seen as trying the reconcile intuitionistic methodology with a realists view on the semantics of mathematics. We will not be able to argue for the realistic point of view or solve the epistemological problems of realism, but hope to formulate a convincing account to the semantics of mathematics. As reason to insist on using realistic mathematics we are content with the fact that constructive mathematics is not able to prove various theorems that seem intuitively valid, and that constructive mathematics might be too weak to carry out all work in the empirical sciences[1].

In liberal constructivism, we want to combine the best of both worlds —

1 See, for example, Hellman (1993a); Hellman (1993b). More recently, Davies (2003) has shown that large parts of quantum mechanics can indeed be carried out in a constructive setting — it seems like an open question if classical reasoning is necessary in the sciences.

to use constructive reasoning when possible, and allow for the use of the law of excluded middle where necessary. Crucially, we want that both reasoning principles are possible in the same system, in contrast to Davies (2005), who has argued that classical and constructive reasoning are both coherent, but should be carried out in different frameworks. A pluralism of frameworks is unsatisfactory for an all-encompassing philosophy of mathematics, since we want a "unity of mathematical reasoning" (Tait, 1983:p. 173). Our proposal is related to Tait's account:

> [...] a critique of the intuitionistic conception of meaning and logic leads, I think, to a promising conception of mathematics on the basis of which these difficulties are resolved and according to which constructive mathematics appears as a part of classical mathematics rather than as a separate science dealing with an entirely disparate subject matter.
>
> — Tait (1983):p. 173

In the account that we will present, classical mathematics rather appears as a part of constructive mathematics: Since we adopt the intuitionistic account to the meaning of the logical constants, classical theorems are those whose proofs lack computational content. We take the formal machinery developed by intuitionism and collapse its constructive character when reasoning classically. It is open for debate if our explanation of what this means for the semantics of mathematics is satisfactory.

1.1 Outline

In Chapter 2, we will present all formal tools that we will use throughout the thesis. We will highlight the juxtaposition of judgements and propositions that is integral to intuitionistic type theory. We will explain the intuitionistic conception of propositions and contrast this with the classical conception of propositions, which can also be reflected in HoTT. Finally, we will get to know the peculiar realization of equality in HoTT.

Afterwards, we will take a closer look at the development of intuitionistic logic in Chapter 3. We will trace the history of intuitionism and see how logic was conceived in this school of thought. We will highlight the difference between formal systems and meaning explanations thereof to understand the relation between the BHK meaning explanations and intuitionistic logic.

Martin-Löf's intuitionistic type theory was initially developed without meaning explanations, only in 1982 we can find a first justification of the system. Chapter 4 is devoted to tracing Martin-Löf's meaning explanations and giving them a modern disguise for HoTT. We will close with highlighting some connections to other developments in proof theory.

Crucially, the informal justification of type theory invalidates the characteristics of identity in intensional type theory, which is at the core of HoTT.

There are several attempts to resolve this problem, we will introduce them in Chapter 5 and see that no attempt is quite satisfactory for explaining the constructive character of identity in HoTT. We will not be able to formulate such a an account as well, but hint at what is required to formulate a satisfying solution.

The final Chapter 6 is devoted to sketching a liberal constructivists philosophy of mathematics. In particular, we will develop a semantical theory that does justice to the importance of constructive proofs while still allowing for the application of the law of excluded middle. We will see which form of the LEM can be used in HoTT and spell out what the assumption of a determinate mathematical universe means for the meaning explanations of intuitionistic type theory.

1.2 Open Problems

The expositions in this thesis seldom commit to one solution of the aforementioned problems, the main contribution should be considered in the fact that these problems are pointed out. The most prevalent problem is that intuitionistic type theory has been developed with an the inferentialistic understanding of the logical constants, while an application of the law of excluded middle seems to ask for a truth-conditional take on semantics. Bridging what could be considered "truthmaker" semantics with truth-conditional semantics could not be carried out in this thesis, but we hope that the groundwork for such a bridge has been laid. A unified semantical theory is necessary if HoTT wants to claim a foundational status and allow for the LEM where necessary. The author is not aware of an issue that would fundamentally impede the development of a coherent theory of meaning for liberal constructivism, but there certainly is an acute need for more explication.

Another issue that we could not resolve is the fact that identity in intensional type theory remains somewhat mysterious. The homotopy interpretation of identity is very fruitful and highly interesting, but falls short of justifying identity as a logical constant. Should we take literal the spatial notions of the homotopy interpretation? Or can we develop a different understanding that does justice to the intricate structure of identity while remaining "logical", in a sense to be spelled out?

Is the present project located in the philosophy of mathematics or in the philosophy of logic? An intuitionist might say that this does not make any difference, since logic is just a very general part of mathematics. Since we are realists and maintain that the mathematical expressions of HoTT refer, whereas the logical expressions do not refer, we face the challenge of drawing a meaningful line between logic and mathematics. The conflation of logical and mathematical notions is what makes intuitionistic type theory so powerful, so we might have to adapt our understanding of what discerns logic from mathematics in this context.

These are only the most blatant problems, HoTT provides a plenitude of highly interesting new ideas that require further investigation. Remarkably, in the debate on whether HoTT can serve as a foundation for mathematics, held, e.g., by Ladyman and Presnell (2016a) and Tsementzis (2017a), its constructive nature has so far not been a matter of concern. A mathematical foundation should allow for classical reasoning where necessary, this thesis can be seen as an assurance that this is unlikely to be an impediment for HoTT becoming an all-encompassing mathematical foundation.

Chapter 2

Judgements and Propositions

The basic entities in intuitionistic type theory are the eponymous *types*. A type is understood as a structurally defined collection that yields "no surprises"[1]: When introducing a type, it should be evident if an element is an inhabitant of a type or not. In contrast to set theory, where the membership of some element in a set needs to be proven explicitly, the membership of some *term* x in a type A, denoted $x : A$, does not require the construction of an intricate proof. Consequently, the statement $x : A$ is not considered a proposition in type theory, but instead called a *judgment*. In particular, we cannot build more complex statements from judgements, for example, $\neg(x : A)$ is a meaningless expression in HoTT. We only state a judgement if it is the case, and we do not give a proof of a judgement since it is *evident*, which we will explicate by requiring that judgements can be decided by a terminating algorithm. Asserting that a judgment is the case is also called a *demonstration*.

Membership of a term in a type is not the only judgment expressible in intuitionistic type theory, another important judgment is the postulation of an equality, aptly called *judgmental equality*. We will introduce all four kinds of judgments devised by Per Martin-Löf in Section 2.1. We will then turn to the understanding of propositions in type theory[2] in Section 2.2, where we will also introduce all types that are used to represent the logical constants. In intuitionistic type theory, propositions and proofs thereof are first-order mathematical objects on the same level as mathematical objects such as the natural numbers. We introduce how the natural numbers are represented in HoTT in Section 2.3. The aforementioned judgmental equality is a weak identity that only works on the level of syntax, identity between

1 This is related to structural definitions most prevalent in category theory. The relevance of these have been brought to the philosophical discourse, e.g., by Awodey (1996) and McLarty (1988).

2 We will call the common core of Martin-Löf's intuitionistic type theory, as presented in Martin-Löf (1975), Martin-Löf (1982) and Martin-Löf (1984), and which is used in HoTT Book (2013), simply "type theory" in the following, ignorant of other type theories such as Russell's ramified theory of types.

mathematical objects can be expressed as a proposition in HoTT by means of introducing a special type. We will turn to the exposition of this identity type in Section 2.4 and highlight the differences between judgmental and propositional equality in Section 2.5. Equipped with this fine-grained handling of equality, we will return to our discussion of propositions and see how we can recover a "classical" understanding of propositions in HoTT in Section 2.6. We will close this chapter with introducing univalence, one of the novel features of HoTT, in Section 2.7.

2.1 Judgements

The basic mathematical statements expressible in Martin-Löf's type theory are *judgments*. In Martin-Löf (1982), we can find the first comprehensive exposition of the four kinds of judgments. If a proposition P is proved to be true, we will have a judgment "P is true", signifying that the proposition has been proved. The idea that the assertion of a proposition is different from a proposition has already been formulated by Arend Heyting:

> Die Behauptung einer Aussage ist selbst wieder nicht eine Aussage.
>
> *(The assertion of a proposition is itself not a proposition.)*
>
> — Heyting (1931):p. 113

Martin-Löf (1987) gives a detailed account of the etymology and philosophical significance of judgments. We will only give an abridged account of the concept here and will deviate from Martin-Löf in a crucial point: If a judgment is valid, it should be *evident* — hence, we do not give proofs for a judgment, but can only require self-explanatory demonstrations. In explaining what "evident" means, Martin-Löf (1987) draws from phenomenological traditions and emphasizes that evidence is a highly subjective notion. Since the subjectivity of intuitionistic mathematics is what we want to depart from in Chapter 6, we want to take another conception of evident statement as basis: Judgments should be *decidable* by a computer. This means that in a finite time, a computer can check if a judgment holds by following a fixed and unequivocal set of instructions. What more can we desire from a criterion of evidence, but that a mindless machine can verify it? In particular, this gives us an objective specification of evidence.

We can make the following four kinds of judgments in intuitionistic type theory:

- A *type* (A is a type)
- $a : A$ (a is a term of type A)
- $A \equiv B$ (A and B are equal types)

- $a \equiv b$ (a and b are equal terms of type A) [3]

The first judgment states that A is a valid type, which corresponds to it being well-formed according to the type formation rules. The second statement corresponds to the statement "A is true", which we can only establish by giving an object a that acts as a proof of A (we will discuss the interpretation of types as propositions in Section 2.2). In the literature on intuitionistic type theory, the notions "element", "term" and "inhabitant" are used more or less interchangeably for entities such as a. In general, we will refrain from using "element" in this context to emphasize the difference to set theory. The word "term" has a distinctively syntactical character, we will often use it due to its conciseness. This should not introduce some juxtaposition to types as less syntactical entities, as types carry the same semantic weight as terms (where the exact weight of both terms and types depends on the underlying meaning theory as we will see in the course of this thesis). The most apt notion we will try to use when doing conceptual work is "inhabitant".[4]

The two latter judgments introduce equalities between types and terms, respectively, which are commonly called *judgemental equalities*. The meaning explanations which we will introduce in Chapter 4 will ensure that judgmental equality for both types and terms is an equivalence relation. In this chapter, we will also make precise how a judgment can be demonstrated. For example, the judgments $s(0) : \mathbb{N}$ and $s(s(0)) \equiv \mathsf{s}(s(0), s(0))$ can be demonstrated automatically.

In the course of building up a mathematical theory, a mathematician will utter more and more judgments. This ordered list of judgments builds up the *context*. For example, we can postulate the following context:

$$A\ type,\ a : A,\ B\ type,\ A \equiv B,\ b \equiv a$$

Note that the above assertions extend the judgmental equality relation: In $b \equiv a$, the definiendum b is defined to be equal to the definiens a, which we already know to be an inhabitant of A. We may sometimes emphasize a definition by writing $b :\equiv a$, but in general, HoTT does not distinguish between definiens and definiendum. For example, both A and B have been introduced before stating $A \equiv B$.

We may use Γ as a variable signifying an arbitrary context. In a context, we may establish that another judgment holds, for example:

$$\Gamma \vdash b : B$$

3 We follow the symbol convention of the HoTT Book (2013). In his original presentation, Per Martin-Löf writes $A = B$ for $A \equiv B$ and $a = b : A$ for $a \equiv b$. Since both terms always belong to the same unique type, we can omit the type information in judgmental equality.

4 In the exploration of HoTT as a foundation for mathematics, Ladyman and Presnell (2016a) have argued for using "token" for inhabitants of a type. We think that the analogy to the philosophy of language is misleading, since under this reading, both 2 and 4 are tokens of the same kind, which suggests that the difference between 2 and 4 is only one of instantiation — but obviously, both expressions have a very different meaning.

If we take the above list of judgments as context Γ, this is a correct assertion. Note that ⊢ is not understood as the derivability predicate as common in logic today. Rather, it is the postulation of an assertion, which also stands in the intuitionistic tradition of Heyting:

> Dass eine Aufgabe gelöst ist, wird angegeben, indem man das Zeichen ⊢ davor setzt; eine Formel, die diese Zeichen enthält, stellt keine Aufgabe mehr dar, sondern eine Mitteilung über die Lösung einer Aufgabe.
>
> (*The solution of a task is indicated by prepending the sign ⊢; a formula containing this sign is not a task anymore, but the announcement of the solution of a task.*)
>
> — Heyting (1934):p. 14

2.2 Propositions

Under the intuitionistic view, a proposition is defined by laying down what counts as a proof of the proposition, and a proposition is true if a proof for it can be given. Hence, truth and provability are conflated (remember that we will consider how a more classical conception of truth can be incorporated in HoTT in Chapter 6). While under the classical conception, a proposition only has a truth value, there are in general many different proofs of a proposition in intuitionistic systems. Consequently, a proposition is identified with the set of its proofs in most constructive formal systems, as we can find for example in Tait:

> There are, at first blush, two kinds of construction involved: constructions of proofs of some proposition and constructions of objects of some type. But I will argue that, from the point of view of foundations of mathematics, there is no difference between the two notions. A proposition may be regarded as a type of object, namely the type of its proofs. Conversely, a type A may be regarded as a proposition, namely the proposition whose proofs are the objects of type A. So a proposition A is true just in case there is an object of type A.
>
> — Tait (1994):p. 51

In Martin-Löf's intuitionistic type theory, some of the proofs of a proposition have a distinct character, namely the canonical proofs. We will take a closer look at the canonical proofs in Chapter 4 and note for the time being that Martin-Löf identifies a proposition with the set of its canonical forms:

> If we take seriously the idea that a proposition is defined by laying down how its canonical proofs are formed (as in the

> second table above) and accept that a set is defined by prescribing how its canonical elements are formed, then it is clear that it would only lead to unnecessary duplication to keep the notions of proposition and set (and the associated notions of proof of a proposition and element of a set) apart. Instead, we simply identify them, that is, treat them as one and the same notion. This is the formulae-as-types (propositions-as-sets) interpretation on which intuitionistic type theory is based.
>
> — Martin-Löf (1984):p. 13

We will get to know how the canonical proofs of the logical constants can be built in Section 2.2.2. Before, we need to become acquainted with dependent types, which can be seen to represent predicates in intuitionistic type theory in Section 2.2.1.

2.2.1 Dependent types

Under the classical view, a predicate maps each object of a collection to a truth value. In intuitionistic type theory, a proposition is not conceived as a truth-apt sentence, but as the set of its proofs. Consequently, a predicate in type theory maps each object of a collection to a collection of objects. More formally, it maps any inhabitant of a type to a distinct type. For example, $\mathsf{Even}(n)$ may express the predicate of being an even number. We can write

$$n : \mathbb{N} \vdash \mathsf{Even}(n)\ \mathit{type},$$

to denote that for any $n : \mathbb{N}$, $\mathsf{Even}(n)$ is a valid type. We may also denote a dependent type as a function $\mathsf{Even} : \mathbb{N} \to \mathcal{U}$, where the function arrow $\to$ and the universe $\mathcal{U}$ will be explained below in Section 2.2.2 and Section 2.7.

Under appropriate definition, $\mathsf{Even}(n)$ will contain at most one inhabitant, namely, when the given n is even. Hence, Even can be seen to just convey a classical proposition as will explained in Section 2.6. In general, a dependent type contains more structure. For example, if we want to represent a date in type theory, we can express the collection of all months with type M. The number of days in a month depends on the respective month, which we can represent with a type dependency:

$$m : \mathsf{M} \vdash d : \mathsf{D}(m)$$

For example, the type $\mathsf{D}(1)$ representing January will contain the inhabitant $31_{\mathsf{D}(1)}$, whereas February, represented by $\mathsf{D}(2)$, does not contain such an inhabitant. Even though both types $\mathsf{D}(1)$ and $\mathsf{D}(2)$ contain a term for the first day of the month, the terms are different: $1_{\mathsf{D}(1)} : \mathsf{D}(1)$ is distinct from $1_{\mathsf{D}(2)} : \mathsf{D}(2)$. Inhabitants of different types cannot be investigated in terms of judgmental or propositional equality, they simply live in different realms.

The intuitionistic reading of predicates can be extended to relations, for example, a dependent type $R : (A \to A) \to \mathcal{U}$ can be considered a reflexive relation if we can prove that for all $x : A$, there is an inhabitant of $R(x, x)$. We will now see how the logical constants are captured in HoTT.

2.2.2 The logical constants in HoTT

In proof-relevant systems, defining the meaning of the logical constants amounts to laying down how proofs of propositions containing the constants can be built. Since proofs and objects are identified, the rules for constructing elements of a type are seen to correspond to the usual rules of inference for propositions in a deductive calculus. In this section, we will introduce the basic types that will represent conjunction, disjunction, implication and universal and existential quantification. We will investigate the exact correspondence between the type formers and the logical constants in Section 3.2 and Section 4.1.

We will prototypically introduce the *dependent function type*, which can be seen to introduce universal quantification in HoTT. The type formation rule specifies how well-formed types look, in case of the dependent function type Π we assume that well-formed types A and $B(x)$ are given to introduce a well-formed dependent function type:

$$\frac{\Gamma \vdash A\ type \quad \Gamma, x : A \vdash B\ type}{\Gamma \vdash \prod_{x:A} B(x)\ type}(\Pi\text{-formation})$$

The introduction rule specifies how inhabitants of a type can be built:

$$\frac{\Gamma, x : A \vdash b : B}{\Gamma \vdash \lambda(x : A).b : \prod_{x:A} B(x)}(\Pi\text{-introduction})$$

An inhabitant of the dependent function type should produce for any given $x : A$ an inhabitant of $B(x)$. When reading $B(x)$ as a proposition, this confirms with the reading of the dependent function type as universal quantification. Note that we have left implicit that $A\ type$ etc., in the following, we will mostly assume that all types are well-formed.

The elimination rule defines how we may use inhabitants defined by the introduction rules. In the case of the dependent function type, we may want to receive the value of $B(x)$ for a given x, where $B[a/x]$ indicates that we have replaced all occurrences of x in B with a[5]:

$$\frac{\Gamma \vdash f : \prod_{x:A} B(x) \quad \Gamma \vdash a : A}{\Gamma \vdash f(a) : B[a/x]}(\Pi\text{-elimination})$$

The computation rules relate the introduction and elimination rule by defining how the elimination rule operates on the inhabitants generated by

5 The formal definition of substitution can be found in HoTT Book (2013):A.2.2.

the introduction rules. For the dependent function type, we will replace any occurrence of x in b by a when applying a function:

$$\frac{\Gamma, x : A \vdash b : B \quad \Gamma \vdash a : A}{\Gamma \vdash (\lambda(x : A).b)(a) \equiv b[a/x] : B[a/x]} \,(\Pi\text{-computation})$$

Note that we have extended the judgmental equality relation with the above rule: Whenever an expression contains a function that is supplied with an argument, we can evaluate the function. This turns judgmental equality from a symmetric relation to something directed, as the evaluation of a function will in general not be reversed. We will see in Chapter 4 how this notion of direction gives rise to view judgmental equality as a reduction relation.

In case type B does not depend on the value of x, we can recover classical functions. A function produces for any $x : A$ an inhabitant of B where B is now a constant type independent of the given x. We can introduce the classical notation for *functions*:

$$A \to B :\equiv \prod_{x:A} B$$

In a similar fashion to the dependent function type, we can introduce the *dependent pair type* to capture existential quantification:

$$\sum_{x:A} B(x)$$

Elements of such a type are pairs (a, b), where $a : A$ can be seen as the witness of the proposition expressed by $B(x)$ and b as the proof that a is indeed a witness, i.e., we have $b : B(a)$. In the following, the scope of the dependent function and pair type will be until the rest of the expression, unless contained with parentheses.

If the type B is constant for any given $x : A$, we can recover the cartesian *product* type from the dependent pair type:

$$A \times B :\equiv \sum_{x:A} B$$

Since B does not depend on A, an inhabitant of $A \times B$ is simply a pair (a, b) with $a : A$ and $b : B$.

The last type missing for an account of the logical constants is one that captures disjunction. The type representing disjunction is commonly called *coproduct*, since in categorical semantics of type theory, it is dual to the product. We write coproducts as follows:

$$A + B$$

Inhabitants of $A + B$ are either of type A or type B, and we want to keep track of which type has been injected. This is done by introducing two distinct introduction rules for $\mathsf{inl}(a) : A + B$ if $a : A$ and $\mathsf{inr}(b) : A + B$ if $b : B$.

We have seen above for the dependent function type how a lambda term $\lambda(x : A).b$ is dispatched with a computation rule. Term constructors that will be dispatched by computation rules are in general called *recursors*. Recursors allow for concise definition of functions since they specify what should happen to any possible inhabitant of a type. In the case of the coproduct, when defining a function that takes as input $A + B$, we need to give a function z_a that specifies what happens if we are given an inhabitant of A, and a function z_b that specifies when given an inhabitant of B. The recursor is then evaluated as follows:

$$R_+(\mathsf{inl}(a), z_a, z_b) :\equiv z_a(a)$$
$$R_+(\mathsf{inr}(b), z_a, z_b) :\equiv z_b(b)$$

We refer to HoTT Book (2013):A.2 for a concise definition of all rules of HoTT.

In addition to types capturing the logical constants, we will also want to represent truth and falsity in the system. Since in intuitionistic type theory, a true proposition is one that is inhabited, falsity is represented with a type that has no constructors:

$$\mathbf{0}$$

The elimination principle for the empty type states that given an inhabitant of $\mathbf{0}$, we can construct an inhabitant of any type. Hence, *ex falso sequitur quodlibet* is a valid inference rule in intuitionistic type theory. As is common in intuitionistic systems, negation is not a primitive logical constant, but introduced by means of implication and negation:

$$\neg A :\equiv A \to \mathbf{0}$$

The dual to the empty type is the unit type:

$$\mathbf{1}$$

The unit type has exactly one inhabitant $\star : \mathbf{1}$, and no other inhabitant. We will explain in Section 2.6 how the unit type can be seen to reflect the traditional understanding of a proposition as a simple truth value.

2.3 Natural Numbers

As we have seen in Section 2.2, proofs are treated just in the same way as mathematical objects. In the intuitionistic tradition, logic and mathematics are deeply intertwined and treated with the same tools. In this section, we will introduce how the prototypical mathematical object, the set of natural numbers, is dealt with in HoTT. For this, we introduce $\mathbb{N}$ *type* and two terms over the natural numbers, representing 0 and the successor function, respectively:

$$\begin{aligned} 0 &: \mathbb{N} \\ \mathsf{s} &: \mathbb{N} \to \mathbb{N} \end{aligned}$$

With these terms we can construct infinitely many inhabitants of the natural numbers, e.g., the term $\mathsf{s}(\mathsf{s}(0))$ represents the natural number 2. In order to define functions over the set of all natural numbers, we introduce the following recursor:

$$\begin{aligned} R_{\mathbb{N}}(0, z_0, z_s) &:\equiv z_0 \\ R_{\mathbb{N}}(\mathsf{s}(n), z_0, z_s) &:\equiv z_s(n, R_{\mathbb{N}}(n, z_0, z_s)) \end{aligned}$$

The recursor for the natural number can be seen to embed induction: Any primitive recursive function over the natural numbers can be defined by stating the result for 0 and the result for $\mathsf{s}(n)$ for an arbitrary n. In order to compute the result of the recursor for a number $\mathsf{s}(n)$, the result of applying the recursor to the predecessor n is computed. This will be iterated recursively until the base case for 0 is computed. For example, addition can be formulated as follows:

$$\mathsf{add} :\equiv R_{\mathbb{N}}(\lambda(n : \mathbb{N}).n, \lambda(n : \mathbb{N}).\lambda(g : \mathbb{N} \to \mathbb{N}).\lambda(m : \mathbb{N}).\mathsf{s}(g(m)))$$

This unfolds in the following natural definition of addition:

$$\begin{aligned} \mathsf{add}(0, n) &\equiv n \\ \mathsf{add}(\mathsf{s}(m), n) &\equiv \mathsf{s}(\mathsf{add}(m, n)) \end{aligned}$$

In this way, all of arithmetic can be formalized in HoTT.

2.4 Propositional Equality

Judgmental equality, which we have introduced in Section 2.1, has a limited power: All judgmental equalities need to be evident in the sense of being demonstrable by a mechanical procedure. Mathematics, however, is full of equality statements that are very much not evident. In order to capture these *propositional equalities*, we will introduce a dedicated type. Under the intuitionistic dogma, a proposition is identified with the collections of its proofs, hence, we will have in general a set of proofs for identity. Since proofs are first-order mathematical objects, we might again investigate the relation among the identity proofs. This gives rise to a behaviour of identity that is highly unfamiliar from a traditional perspective, which will occupy us in Chapter 5.

We have the following type representing identity between two inhabitants a, b of type A:

$$\mathsf{Id}_A(a, b)$$

There is only one term constructor, which states that any element $x : A$ is identical to itself:

$$\mathsf{refl}_x : \mathsf{Id}_A(x, x)$$

On a first reading, this seems to suggest that there is only one inhabitant of an identity type. However, the *family* of identity types is defined by the term constructor above, and since an identity type depends on a, b and A, there are in general multiple inhabitants of the identity type. This means that we might have a proof of identity $p : \mathsf{Id}_A(a, b)$ different from both refl_a and refl_b. Since proofs are again mathematical objects, we can investigate the equality between p and refl_a by means of the following type:

$$\mathsf{Id}_{\mathsf{Id}_A(a,b)}(p, \mathsf{refl}_a)$$

For example, if we consider the integers modulo 2, $\mathbb{Z}/2\mathbb{Z}$, as a groupoid, we have non-trivial identity proofs: We have a proof $p : \mathsf{Id}_{\mathbb{Z}/2\mathbb{Z}}(0, 2)$, but p is not judgementally equal to both refl_0 and refl_2. However, we can identify all proofs of equality on a propositional level for this specific type, e.g., we will be able to provide an inhabitant of $\mathsf{Id}_{\mathsf{Id}_{\mathbb{Z}/2\mathbb{Z}}(0,2)}(p, \mathsf{refl}_0)$.

In general, proofs of identity do not need to be propositionally equal. This means that investigating identity between identity proofs can be continued indefinitely which yields a very rich structure known in category theory as an ∞-groupoid and in algebraic topology as a homotopy type.

Just like the elimination rules for the logical constants sketched in Section 2.2.2, there is an elimination rule for the identity type, commonly referred to as the *J-rule*. The J-rule specifies how an identification $p : \mathsf{Id}_A(x, y)$ of two elements $x, y : A$ may be used. If we have proven that an arbitrary dependent type $C(x)$ is inhabited for x, we expect that it is also inhabited for y, since both terms are equal. In other words, we want that any type C respects propositional equality. In general, the type C might not only depend on x, but also refer to y and the identification p between both terms.

More precisely, the J-rule states that we only need to provide an element of C for element x and the trivial self-identification:

$$c : \prod_{x:A} C(x, x, \mathsf{refl}_x)$$

Then the J-rule ensures that $C(x, y, p)$ holds for all x, y and proofs of their equality p:

$$f : \prod_{x,y:A} \prod_{p:\mathsf{Id}_A(x,y)} C(x, y, p)$$

The function f will yield the given proof c when only x is provided, i.e., we have the computation rule $f(x, x, \mathsf{refl}_x) :\equiv c(x)$. The homotopical reading of HoTT interprets the J-rule as *path induction* in a homotopy space, we will explain this interpretation in Section 5.2.

When applying the J-rule, the type $C(x, y, p)$ is sometimes called the *motive*. For example, if we take as motive the type $P(x) \to P(y)$ for an arbitrary dependent type $P : A \to \mathcal{U}$, we can prove indiscernibility of identicals as follows: We can construct a function $c : \prod_{x:A} P(x) \to P(x)$ with the identity function:

$$c :\equiv \lambda(x : A).\mathsf{id}_{P(x)}$$

The J-rule then gives rise to the following function:

$$f : \prod_{x,y:A} \prod_{p:\mathsf{Id}_A(x,y)} P(x) \to P(y)$$

Under the propositions-as-types interpretation, the function f expresses the indiscernibility of identicals: Whenever we have a proof of x and y being equal and know that $P(x)$ holds, we also know that $P(y)$ holds. Topologically, this construction is known as *transport*.

One can quickly show that the J-rule ensures that propositional equality is symmetric and transitive (HoTT Book, 2013:Lemma 2.1.1, Lemma 2.1.2), reflexivity follows trivially from the existence of refl_x for any $x : A$. Hence, propositional equality is an equivalence relation as expected.

Of interest in the course of this thesis will be a variant of the J-rule which is equivalent to the rule above. This variant is termed *based path induction* in the HoTT Book (2013), since one of the endpoints of the identity path is fixed. We formulate identity elimination for a fixed $a : A$ and variable $x : A$. This means that given a type family:

$$\mathsf{C} : \prod_{x:A} \mathsf{Id}_A(a, x) \to \mathcal{U}$$

We need to construct a proof of C for the fixed element a and the trivial self-identification:

$$c : \mathsf{C}(a, \mathsf{refl}_a),$$

Based path induction then yields a function that produces a witness of C for any $x : A$ that can be identified with a:

$$f : \prod_{x:A} \prod_{p:\mathsf{Id}_A(a,x)} C(x, p)$$

In particular, if we apply f to the base point a, we will just retrieve the constructed proof c of C for a, i.e., $f(a, \mathsf{refl}_a) :\equiv c$. Based path induction is equivalent to path induction stated above (HoTT Book, 2013:Sect. 1.12.2).

2.5 Equality, Revisited

We have seen two kinds of equality, judgmental equality in Section 2.1 and propositional equality in Section 2.4. We want to further compare both notions of equality in this section. Propositional equality has been characterized as "internal", in contrast to judgmental equality as an "external" equality (Ladyman and Presnell, 2014). Indeed, propositional equality can be seen as identifying mathematical objects, whereas judgmental equality is generated by how the mathematician sets up the system. In more traditional terms, we can regard judgmental equality as equality of sense, and propositional equality as equality of reference. The idea of associating the different equalities with a Fregean distinction between intension/sense/meaning and extension/reference/value can be already found in Martin-Löf's expositions:

> We then have four kinds of equality:
>
> (1) $\equiv$ or $=_{def.}$,
>
> (2) $A = B$,
>
> (3) $a = b \in A$,
>
> (4) $I(A, a, b)$.
>
> Equality between objects is expressed in a judgement and must be defined separately for each category, like the category sets, as in (2), or the category of elements of a set, as in (3); (4) is a proposition, whereas (1) is a mere stipulation, a relation between linguistic expressions. Note however that $I(A, a, b)$ true is a judgement, which will turn out to be equivalent to $a = b \in A$ (which is not to say that it has the same sense). (1) is intensional (sameness of meaning), while (2), (3) and (4) are extensional (equality between objects). As for Frege, elements a, b may have different meanings, or be different methods, but have the same value. For instance, we certainly have $2^2 = 2 + 2 \in N$, but not $2^2 \equiv 2 + 2$.
>
> — Martin-Löf (1984):p. 59–60

However, in our presentation, which aligns with the HoTT Book (2013), we conflate equalities (1) with (2) as judgmental equality between types and (1) with (3) as judgmental equality between terms. Common to all these equalities is that they solely rely on the comparison of symbols. Hence, we can carry out a mechanic reduction process to decide $\equiv$ for both terms and types, as we will see in Chapter 4. As Martin-Löf points out, $I(A, a, b)$ (in our notation $\mathsf{Id}_A(a, b)$), is a proposition, whereas the assertion of a proposition is a judgment, as discussed in Section 2.1.

From the modern perspective of HoTT, we have to revise the example given by Martin-Löf in the quote above: We have $2^2 \equiv 2 + 2$ as a judgmental equality. This is due to the fact that when unfolding the definition of exponentiation and addition, both sides of the equation can be reduced to the same syntactical expression, namely, the term representing 4. Arithmetical statements with fixed numerals can always be decided. With this updated knowledge of what counts as an evident equality, we can fully subscribe to the following characterization of the two equalities by Martin-Löf:

> Definitional equality $\equiv$ is a relation between linguistic expressions; it should not be confused with equality between objects (sets, elements of a set etc.) which we denote $=$. Definitional equality is the equivalence relation generated by abbreviatory definitions, changes of bound variables and the principle of substituting equals for equals. Therefore it is decidable, but not in the sense that $a \equiv b \vee \neg(a \equiv b)$ holds, simply because $a \equiv b$ is not a proposition in the sense of the present theory.
>
> — Martin-Löf (1984):p. 31

Since we conflate definitional equality and the other intensional equalities, we use *judgmental equality*, *definitional equality* and *intensional equality* interchangeably. Judgmental equality implies propositional equality, e.g., if we have $a \equiv b$ for $a, b : A$, then we have $\mathsf{Id}_A(a, b)$ witnessed by refl_a. This goes well with the reading of $\equiv$ being the equality of sense, since if the sense of two expressions is the same, the reference of both expressions certainly will have to be the same, as well.

Note that in general, a statement about arbitrary natural numbers cannot be proven by means of judgemental equality. For example, commutativity of addition needs to be proven by providing an inhabitant of the following type:

$$\prod_{m,n} \mathsf{Id}_{\mathbb{N}}(m + n, n + m)$$

In conclusion, propositional equality statements can be seen as those statements requiring creative proof work, whereas judgmental equality is a syntactical equality that can be taken care off by a machine. This provides a neat demarcation between meaningful equalities of the form "$a = b$", and evident equalities of the form "$a = a$". Frege posed the question of why we are interested in equality statements if they just express something as trivial as identity: Equalities that hold *by definition* are indeed trivial, but identifying different presentations of the same structure is everything but trivial and a worthwhile mathematical endeavour.

2.6 Mere Propositions and Propositional Truncation

Under the classical view, a proposition is a truth-apt sentence. In general, types in HoTT are not only truth-apt, but might contain a plenitude of inhabitants. However, a class of types does resemble classically conceived proposition: Types with at most one inhabitant. This is captured by the following type, which characterizes all types that are *mere propositions*:

$$\mathsf{isProp}(\mathsf{A}) :\equiv \prod_{x,y:A} \mathsf{Id}_A(x,y)$$

Mere propositions have no inner structure since the identity type over x and y is trivial. The collapse of the identity type is upwards-closed, i.e., the identities between identity proofs of inhabitants of a mere proposition are also trivial.

We can transform any type into a mere proposition, this transformation is called *propositional truncation*. The truncated version of a type is denoted:

$$\|A\|$$

$\|A\|$ is inhabited if and only if the original type A is inhabited, and contains no other inhabitants. Formally, every $x : A$ is mapped into an element $|x| : \|A\|$. For any $x, y : A$ we have $\mathsf{Id}_{\|A\|}(|x|, |y|)$. Thereby, $\|A\|$ contains exactly one inhabitant iff A is inhabited, and no inhabitants otherwise.

Note that propositional truncation can be generalized to truncation of the identity type at any level. If we require equality between two inhabitants of a type to be a mere proposition, we recover classical *sets*. This generalizes to *homotopy n-types*, which are introduced in HoTT Book (2013):Ch. 7. Interestingly, truncation at any level induces a different extensional equality: If we truncate all types to mere propositions, propositional equality amounts to logical equivalence. Truncating all types to sets means that propositional equality captures equinumerousity.

2.7 Universes and Univalence

Above, we have only stated *A type* if A was a well-formed type. This was a simplification since in HoTT, every type lives in a universe. More specifically, we have an infinite sequence of universes:

$$\mathcal{U}_0 : \mathcal{U}_1 : \mathcal{U}_2 : ...$$

Every type lives in one $\mathcal{U}_i$. The universes are "cumulative" in the sense that if $A : \mathcal{U}_i$, then also $A : \mathcal{U}_j$ for all $j > i$. This means that every type is itself inhabitant of infinitely many universes, in contrast to terms, which always belong to one unique type. This should be seen as a primarily technical feature of HoTT without great conceptual significance. In this work, we can be content with leaving the index implicit and always writing $A : \mathcal{U}$.

The universe can be seen as a "big" type in the sense of big categories in category theory. In particular, HoTT is strictly predicative: There is no universe containing itself. Consequently, there can be no function $\lambda(i : \mathbb{N}).\mathcal{U}_i$, since we have no universe that could encompass such a function containing all types (and itself in particular). The indices i are different from the type of the natural numbers and can be seen as external to the mathematical objects represented in HoTT.

A crucial new insight presented in the HoTT Book (2013) is the principle of *univalence*, which Vladimir Voevodsky devised after concepts from homotopy theory. There has been considerable interest in the philosophical repercussions of univalence, see for example Awodey (2014), Awodey (2018), Ladyman and Presnell (2016b) and Tsementzis (2016). Univalence can be seen to give formal justification to the dogma of mathematical structuralism: Isomorphic objects can be identified. For example, type theory cannot distinguish between different isomorphic presentations of the natural numbers such that problems in the style of Benacerraf (1965) cannot arise. Unfortunately, we can only touch on the relevance of univalence for our undertakings in Chapter 5, and will only give a technical exposition here. A self-contained technical presentation of univalence can also be found in Escardó (2018).

As a first intuition, univalence can be seen to state that "isomorphic objects might be identified". The notion of isomorphism is generalized to the notion of *homotopy equivalence* which is defined as follows:

$$(A \simeq B) :\equiv \sum_{f:A\to B} \sum_{g,h:B\to A} (f \circ g \sim \mathsf{id}_B) \times (h \circ f \sim \mathsf{id}_A)$$

The notion $\sim$ used above can be seen to capture equality between functions, stemming from homotopy theory: Two functions are considered *homotopic* if they can be continuously transformed into each other. For this, let $f, g : \prod_{x:A} P(x)$ be two functions over the dependent type $P : A \to \mathcal{U}$. Both functions are then considered homotopic, $f \sim g$, if the following holds:

$$(f \sim g) :\equiv \prod_{x:A} \mathsf{Id}_{P(x)}(f(x), g(x))$$

The inhabitant of $f \sim g$ can be seen to be the continuous transformation, since for any given x it shows how f and g can be identified.

For A and B being sets as defined above in Section 2.6, homotopy equivalence boils down to the traditional definition of isomorphism. The more cumbersome formulation above is used to solve some coherence problems for types with more structure than sets, as there might be several ways to prove an isomorphism between types more complex than sets. The formulation of $\simeq$ ensures that there is at most one witness of an equivalence.

In all flavours of intuitionistic type theory, propositional equality implies equivalence, i.e., we can construct the following function:

$$\mathsf{idtoequiv} : \mathsf{Id}_{\mathcal{U}}(A, B) \to (A \simeq B)$$

The univalence axiom states that the converse also holds, i.e., that we can conclude from the equivalence of two types that both types are equal:

$$\mathsf{ua} : (A \simeq B) \to \mathsf{Id}_{\mathcal{U}}(A, B)$$

In particular, both functions induce an equivalence between equivalence and identity, which leads to the following formulation of *univalence*:

$$\mathsf{Id}_{\mathcal{U}}(A, B) \simeq (A \simeq B)$$

A universe $\mathcal{U}$ in which the above type is inhabited is said to be *univalent*, i.e., univalence is a property of universes.

Univalence can be seen to establish a one-to-one mapping between identity and equivalence — consequently, the identity type has to be as rich in structure as equivalence between types. This means that there need to be identity proofs different from trivial self-identifications to capture this rich structure. Awodey concludes:

> Rather than viewing it [the univalence axiom] as identifying equivalent objects, and thus collapsing distinct objects, it is more useful to regard it as expanding the notion of identity to that of equivalence.
>
> — Awodey (2014):p. 9

Crucially, it is not possible to construct a term ua in the version of type theory presented in HoTT Book (2013). There is plentiful research effort on devising a type theory were univalence is given a constructive character, i.e., where it is possible to construct a function ua. There are very auspicious *cubical type theories* in which univalence holds as a theorem.

Chapter 3

Constructive Logic

In the upcoming chapters, we will investigate Martin-Löf's meaning explanations and the purported view on the constructive content of proofs. To put Martin-Löf's work in context, we will sketch in this chapter the historic development leading to the development of intuitionistic type theory. As the name suggests, the development of intuitionistic mathematics and logic provides the background against we can make sense of Martin-Löf's meaning explanations. We will shortly look at the genesis of intuitionism with the works of Brouwer in Section 3.1 and turn to the first systematic formalization of intuitionistic reasoning in Section 3.2, commonly referred to as the BHK interpretation. We will see how the BHK account can be made formally precise in Section 3.3, before sketching in Section 3.4 the works of Bishop, which called Martin-Löf's attention to the need of formulating a more powerful language for formalizing constructive mathematics.

3.1 Brouwer and the Advent of Intuitionism

The idea that a proof of an existential assertion has to be accompanied with a construction of the respective object goes back to Kronecker in the 19th century. The main developer for and advocate of intuitionism in the *Grundlagenstreit*, Luitzen Brouwer, took this idea and formulated a systematic criticism of mathematical practice at that time. Under his view, talking of a mathematical universe that is mind-independent was metaphysical speculation that should be stripped from mathematics. Instead, he treated mathematics as a product of the mind, and mathematical objects as nothing but mental processes. Since mathematical objects do not exist independently of a mind, but are developed incessantly, statements about those objects are not determinately true or false. Instead, the truth of a mathematical statement needs to be "experienced" by a mathematical subject. Brouwer traced his skepticism towards the use of the law of excluded middle back to Immanuel Kant. Based on the Kantian view of time as a sequence of moments, Brouwer developed the *first act of intuitionism* to construct the natural numbers. The *second act of intuitionism* recovers the continuum in accordance with the infinite divisibility of moments. Based on this idea, free choice sequences give rise to the real numbers.

The apparatus developed by Brouwer contradicts parts of classical mathematics, for example, every total function on the continuum is continuous in his framework. In other cases, classically valid theorems cannot be proven constructively. For example, there cannot be an algorithm that produces for a continuous function f on the closed interval $[0, 1]$ with $f(0) < 0$ and $f(1) > 0$ the point x such that $f(x) = 0$. Even though the lemma seems intuitively true and can be proven classically, the representation of the real numbers as only potentially infinitesimal hinders a constructive proof of the intermediate value theorem.

Brouwer rejected the idea of developing a core language for intuitionistic mathematics and sustained a rather idiosyncratic style of reasoning. In his view, natural or logical language may approximate the mental constructions that have been carried out by a mathematical subject, and using language is obviously necessary to work with other subjects, but language cannot be identified with the mental constructions that are mathematics. It is irony of history that the originator of today's framework for formalizing mathematics in a computer was deeply skeptical towards the formalization of mathematics in general.

3.2 Heyting and Kolmogorov, and the Formalization of Intuitionism

It was only Brouwer's student Arend Heyting who set out to develop a formal system that could capture the reasoning principles of intuitionism and specified what should be considered an intuitionistically valid proof. Independently of Heyting, Andrey Kolmogorov developed his own formalization of Brouwer's reasoning principles. Later, Heyting and Kolmogorov agreed that their characterizations were essentially equivalent.

Under the impact of Brouwer, Heyting downplayed the role of intuitionistic logic for intuitionism. According to him, intuitionistic logic was merely a tool, but not a foundation for mathematics since logic itself lacks a foundation. In an imagined debate about the mathematical foundations, Heyting lets the prototypical intuitionist maintain:

> The process by which [a logical theorem] is deduced shows us that it does not differ essentially from mathematical theorems; it is only more general, e.g. in the same sense that "addition of integers is commutative" is a more general statement than "2 + 3 = 3 + 2". This is the case for every logical theorem: it is but a mathematical theorem of extreme generality; that is to say,

> logic is a part of mathematics, and can by no means serve as a foundation for it.
>
> — Heyting (1956):p. 6

Hence, Heyting subscribed to Brouwer's conviction that mental constructions of mathematical objects are prior to logic, and the logical laws just happen to be common to many mathematical statements. The project of Kreisel, the father of modern proof theory, is portrayed by Sundholm as follows:

> We may compare his [Kreisel's] programme with a common interpretation of the logicist programme:
>
> (1) to define mathematical concepts in terms of logic, and
>
> (2) to derive the mathematical theorems as truth of logic.
>
> The parallel is obvious:
>
> (1) to define the logical constants in terms of constructions, and
>
> (2) to derive the truths of logic as theorems of the theory of constructions.
>
> — Sundholm (1983):p. 156

This can be seen to take serious Heyting's project of reversing the logicist programme: Logic is not fundamental, instead the mental constructions are basic and logic can be recovered from these mental constructions.

Heyting referred to a proposition as an "expectation" an "intention" of being proved, Kolmogorov saw a proposition as a "problem" or "task" that needs to be solved. Consequently, a logical constant such a disjunction $A \vee B$ expresses the expectation that either a proof of A or a proof of B is given. This extends to what is commonly referred to as the BHK interpretation. Troelstra and Dalen (1988):p. 9 recapitulate this definition of the logical constants as follows:

- Absurdity $\perp$ has no proof.
- A proof of $P \wedge Q$ is given by presenting a proof of P and a proof of Q.
- A proof of $P \vee Q$ is given by presenting either a proof of P or a proof of Q.[1]
- A proof of $P \to Q$ is a construction which permits us to transform any proof of P into a proof of Q.

1 Troelstra and Dalen (1988) omit in this representation that usually, a proof of $P \vee Q$ also is required to evince which of the disjuncts has been proved.

- A proof of $\forall x.Q(x)$ is a construction which transforms a proof of $d \in D$ into a proof of $P(d)$.
- A proof of $\exists x.Q(x)$ is given by providing $d \in D$, and a proof of $Q(d)$.

The clauses can be read as biconditionals: For example, if we have proved that $P \wedge Q$ holds, we can obtain a proof of P and a proof Q and, vice versa, given a proof of P and Q, we can construct a proof of $P \wedge Q$.

This account makes apparent the conflation of proofs and objects in the intuitionistic tradition: In the clauses for the quantifiers, D is the domain that variables x range over. Giving a proof is then nothing but giving an object of the domain D. Other logicians maintained a difference between objects and proofs in their constructive systems, e.g., we can find in Scott's "Constructive Validity", which can be considered a predecessor of Martin-Löf's type theory:

> The construction is an object of the theory while the proof is an elementary argument about the theory.
>
> — Scott (1970):p. 261

However, Scott is not able to maintain the distinction between proofs and objects and concludes in the Postscript of his paper that "the attempt to eliminate 'proofs' (as abstract objects) and to concentrate on 'pure' constructions is not successful" (Scott, 1970:p. 272). We will return to the ontological character that intuitionistic theories attribute to proofs in Section 4.4 and Chapter 6.

A negated formula $\neg P$ is not defined primitively in the BHK account, instead it is represented with P implying absurdity. We could add a clause for negation as follows:

- A proof of $\neg P$ is a construction which transforms any hypothetical proof of P into a proof of a $\bot$.

It should be noted that the above characterization of intuitionistic reasoning was not without controversy, for example, some intuitionists like Griss did not accept that negation is a sensible mental construction at all (Heyting, 1956:p. 10).

This account to intuitionistic mathematics is informal and uses the notions "proof" and "construction" pre-theoretically without further explication. The term *semantics* is hence not an adequate description of the BHK account, if one understands semantics as a translation from one formal language to another formal language. Rather, the term *meaning explanations* seems like an adequate term for the BHK account to constructive reasoning. The informal account helps to understand the meaning of the logical constants, but does not give a straightforward recipe to reason intuitionistically. Semantics, in contrast, pin down the exact behaviour of a system. Often, semantics are articulated in set theory, such as Kripke's semantics for intuitionistic logic, but we will use the term in general for translation into any formal language. The differentiation between semantics and meaning explanations will be of use for us in the following, and is in accordance with current philosophical nomenclature, e.g., by Atten (2017).

We will now get to know a semantics which gives a clear account of the informal notions of the BHK meaning explanations.

3.3 The Lambda Calculus and Propositions-as-types

One way to make the BHK meaning explanations formally precise is based on the simply typed lambda calculus. The development of the lambda calculus goes back to the 1930s, where Alonzo Church developed it to provide a syntactical model of computation. The discovery that the lambda calculus can be used as a semantics for the implicative fragment of BHK was made by Haskell Curry (1934).

The fundamental insight of Curry is that the types of the lambda calculus can be regarded as the propositions of intuitionistic logic, and terms of a type as proofs of the associated proposition. This allows for identifying a proposition with the set of its proofs, instead of just a truth value. If one wants to maintain the classical view of a proposition as a truth-apt sentence, the correspondence can also be formulated by identifying each type with the *proposition that the type is inhabited*.

The *types* are formed according to the following rules[2]:

$$\tau ::= \tau \to \tau \mid t$$

where t is a an element of T, a set of atomic types.

The *terms* of the simply typed lambda calculus can be formed according to the following rules:

$$e ::= x \mid (x)e \mid e(e) \mid c$$

where x is a variable from a set of infinitely many variables $x, y, \ldots$, $(x)e$ is function abstraction over a variable x (which is implicitly assigned some type τ), $e(e)$ denotes function application, and c is a constant from a given set of constants C. The type assignment is defined as follows: If x is of type τ and e of type τ', then $(x)e$ is of type $\tau \to \tau'$. The term $e(e')$ can only be formed if e is of type $\tau \to \tau'$ and e' is of type τ, the type of $e(e')$ is then τ'.

There is one computation rule specifying how function abstraction and application interact, which is commonly called β-reduction:

$$((x)e)a \longrightarrow e[a/x]$$

where $e[a/x]$ denotes the result of substituting all occurrences of x in e by a. Since x and a have the same type if the expression $((x)e)a$ is well typed, the reduction step preserves the type of the expression.

2 We will use the Backus-Naur form for concise grammar definitions in the following.

With this formal system we can give an explication of the implicative fragment of the BHK account of intuitionistic logic: For the propositions P and Q we introduce atomic types A and B. A proposition $P \to Q$ will then be represented with the type $A \to B$. In order to prove the proposition $P \to Q$, we have to construct a closed term $(x)e$, where x is of type A and e is of type B. A "construction" is hence explicated with the notion of a function, and "transformation" is interpreted as computation of a function.

For example, to prove the proposition $P \to (Q \to P)$, we can construct a term $(x)((y)x)$ of type $A \to (B \to A)$. Since this function is closed, we have proved that the proposition $A \to (B \to A)$ is a tautology.

Note that abstraction is commonly denoted $\lambda x.e$, where sometimes the type of x is stated explicitly, i.e., $\lambda(x : \tau).e$. We will make use of our definition of the lambda calculus later in Section 4.2, where the absence of any special characters in the calculus will come in handy to distinguish the lambda calculus from HoTT.

Howard (1980) extended the idea of Curry to capture predicates and to formalize the full account of BHK. We will present a part of Howard's account in the modern disguise of HoTT in Section 4.1.

3.4 Bishop's Constructive Mathematics

The formalization of the BHK conception of intuitionistic logic allowed for some useful applications, e.g., Heyting's account to arithmetic seemed promising since a large part of the theorems of Peano arithmetic could be carried over to the intuitionistic system. However, in other fields such as analysis the proofs turned out to require significantly more work, if they could be carried out constructively in the first place. The evident success of classically pursued mathematics in the decades following the Grundlagenstreit made the reservations of Brouwer seem like unnecessary philosophical concerns. The intuitionistic project had stalled.

Only in 1967, Errett Bishop drew the attention of the mathematical community back to constructive mathematics. In *Foundations of constructive analysis*, he reconstructed large parts of analysis solely by constructive means. Bishop agreed with Brouwer's criticism of the classical mathematicians disregard of the importance of constructive proofs, but did not adhere to Brouwer's philosophy of mathematics: "There are no dogmas to which we must conform" (Bishop, 2012:p. ix).

Bishop divides mathematics in a "realistic" and an "idealistic" part, where the realistic part corresponds to the constructively proven theorems. The law of excluded middle is an idealistic assumption that is consistent with the realistic body of theorems. Alternatively, the realistic part can be extended to capture the intuitionistic body of mathematics, such that, for example, all total functions on the continuum are continuous.

In *Foundations of constructive analysis*, Bishop pursues an informal style that allows for a concise and comprehensive presentation of constructive proofs. Bishop also worked on transforming his proofs into a semi-formal and eventually completely formal language to be entered into a computer. As Petrakis (2018) has discovered, Bishop gave some sort of dependent type theory already in 1968, but did not publish his system. The informal constructive mathematics pursued by Bishop was going well beyond of what could be captured by intuitionistic logic in the spirit of BHK. There have been multiple attempts at formalizing the semi-formal exposition of Brouwer: For example, John Myhill developed with Peter Aczel and Harvey Friedman intuitionistic set theory, which is based on classical set theory, but changes the logic and axioms of ZFC. Important for us, Martin-Löf also set out to formalize constructive mathematics in the style of Bishop:

> The theory of types with which we shall be concerned is intended to be a full scale system for formalizing intuitionistic mathematics as developed, for example, in the book by Bishop.
>
> — Martin-Löf (1975):p. 73

Chapter 4

Computational Content

After we have seen the conception of constructive logic in the works of the early intuitionists in the previous Chapter 3, we will now turn to the modern development of constructive type theory. Per Martin-Löf was well aware of the correspondence between the lambda calculus and logical formulae discovered by Curry and Howard, and was the first one to directly link the lambda calculus with natural deduction proofs (Sundholm, 2012). Afterwards, he set out to develop systems in the spirit of the lambda calculus, but with more expressive power than first-order logic. The work of Bishop (1967) gave new life to the project of constructive mathematics and subsequently, Martin-Löf set out to provide a formal system for constructive mathematics done in the style of Bishop. Bishop had drawn the attention of the mathematical community to the pragmatic content of proofs, or, as we will call it in the following, the *computational content* of proofs. The level of importance paid to constructive proofs changes from field to field, number theorists for example pay special attention to what they call "effective" proofs: Alan Baker received the Fields medal in 1970 for carrying out an effective proof of a theorem which had already been proved by Klaus Roth in 1958. However, the development of constructive foundations has been mostly unnoticed by number theorists and it seems that the prevalent view on constructive reasoning still is that it forbids use of the law of excluded middle and is very cumbersome. As we will see in this chapter, this is an insufficient description of Martin-Löf's type theory: The theory has expressive power going well beyond intuitionistic logic, and the BHK interpretation only explains some aspects of constructive type theory — we will see in Section 4.1 that a fragment of HoTT can be seen to internalize intuitionistic first-order logic, namely the fragment of HoTT where all types are propositionally truncated. Crucially, the BHK meaning explanations are unable to motivate the untruncated version of HoTT. In HoTT, a type can have multiple proofs, and the proofs themselves can be investigated again. Hence, intuitionistic type theory gives a more fine-grained view on the computational content of proofs and we need a refinement of the BHK conception of constructive reasoning to explain the features of intuitionistic type theory. Per Martin-Löf presented his meaning explanations as such a refinement, in which the computational content of proofs is pinned down by singling out *canonical forms* and explaining the meaning of non-canonical types and terms by reducing them to canonical forms. We will retrace his meaning explanations in Section 4.2.

Afterwards, we will introduce an interesting result for homotopy type theory, namely, *homotopy canonicity*: The computational content of some proofs can be recovered, even if they are obtained by non-constructive means. This refined view on the computational content of proofs will be relevant in our discussion in Chapter 6.

The meaning explanations developed by Per Martin-Löf loosely stand in the Wittgensteinian tradition of determining the meaning of an expression by its use, which has been made precise by Gentzen in proof theory: The introduction rules of his calculus *determine the meaning* of the logical constants (and elimination rules are just corollaries of the introduction rules). There are important parallels of Martin-Löf's meaning explanations and the ideas of proof-theoretic semantics. We will look at these connections in Section 4.4 so we can make use of them in our development of a meaning theory for HoTT in Chapter 6.

In this chapter, we will completely ignore identity. Martin-Löf's explanation of identity was not able to explain its features in intuitionistic type theory, we will take a closer look at the defects of Martin-Löf's meaning explanations and investigate how we could make sense of identity in the context of homotopy type theory in the next Chapter 5.

4.1 BHK in Homotopy Type Theory

In Section 3.3, we saw how the simply typed lambda calculus can be used to make precise the BHK meaning explanations. For that purpose, propositions are interpreted as types in the calculus and proofs as terms of the respective type. To be precise, every type corresponds to the proposition that the type is inhabited. A truncated version of dependent type theory can also be seen to model BHK, as discovered and spelled out by Pfenning (2001) and Awodey and Bauer (2004). The latter work introduces bracket types for erasing the computational content of a type in a dependent type theory. In the HoTT Book (2013), type truncation is used to the same effect, which we introduced in Section 2.6. With truncation, we can make all types to *mere propositions*, i.e., to types with at most one inhabitant. The correspondence between truncated type theory and BHK has sometimes been coined *propositions-as-some-types* since all propositions are expressed by a corresponding type, but not all types express a proposition. We will prototypically present how three clauses of the BHK meaning explanations can be made precise in truncated HoTT.

We can explicate absurdity by representing the proposition $\perp$ with the type $\mathbf{0} : \mathcal{U}$. Since there is no constructor of $\mathbf{0}$, we can formulate the following clause to capture absurdity:

- Absurdity $\mathbf{0}$ has no proof.

In order to represent the other clauses, we will consider arbitrary propositions P and $Q(x)$ to be represented with types $A : \mathcal{U}$ and $B(x) : A \to \mathcal{U}$. We assume A and $B(x)$ to be propositionally truncated for any x, hence, A and $B(x)$ have at most one inhabitant. We can explicate universal quantification with the dependent function type. The introduction rule for the dependent function type states that if we have a function $\lambda(x : A).y$ that takes any element $x : A$ to a $y : B(x)$, then we have constructed a term of type $\prod_{x:A} B(x)$. Hence, we can represent universal quantification with the following clause to make the BHK interpretation precise:

- A proof of $\forall x.Q(x)$ is a function $\lambda(x : A).y : \prod_{x:A} B(x)$ which maps any inhabitant x of A into an inhabitant y of $B(x)$.

We can recover simple functions from the dependent function type $\prod_{x:A} B(x)$ if $B(x)$ does not depend on x:

- A proof of $P \to Q$ is is a function $\lambda(x : A).y : \prod_{x:A} B$ which maps any inhabitant x of A into an inhabitant y of B.

The extension to the other type formers and clauses of BHK is straightforward. Even though this explication of the BHK interpretation resembles the one in the lambda calculus presented in Section 3.3, there is an important difference: The line between types and terms is blurry in dependent type theories, since terms may be used in the definition of a type. We conclude that BHK can be made precise in truncated HoTT and conversely, that the clauses of BHK can be seen as an adequate meaning explanation for the fragment of homotopy type theory were all types are mere propositions, i.e., where all computational content is stripped from a type. This raises the question:

How can we understand the full system of homotopy type theory?

4.2 Martin-Löf's Meaning Explanations

Univalence implies the existence of a type with non-trivial identity proofs, hence, univalence is inconsistent with the requirement that all types are mere propositions. As we saw previously, the BHK meaning explanations can only explain the meaning of some types, namely the truncated ones. However, HoTT has strictly more expressive power then intuitionistic first-order logic since we can refer to all proofs as objects in the system itself and investigate the relation between those proof objects.

Per Martin-Löf introduced the meaning explanations for intuitionistic type theory in Martin-Löf (1982) and Martin-Löf (1984) to justify the constructive validity of his system. He wanted to show that the rules of inference are *evident* (in a sense to be explained). The meaning explanations of Per Martin-Löf are not only informal intuitions for his type theory, but can be stated in terms of a *semantics*, namely by translating all terms and types of intuitionistic type theory into a simply typed lambda calculus. This approach has recently been given a concise presentation by Dybjer (2012), we will extend this approach and follow the conviction that a formal model of the meaning explanations will help us to make our intuitions precise:

> It is sometimes said that the meaning explanations are nothing but a realizability interpretation (in the sense of Kleene), but this is fundamentally misleading. Realizability provides a meta-mathematical and not a pre-mathematical interpretation! Nevertheless, it helps us to be precise and to understand the details involved in the meaning explanations.
>
> — Dybjer (2012):p. 223

Notably, the meaning explanations presented in Martin-Löf (1982) and Martin-Löf (1984) invalidate intensional type theory, which is the bedrock of homotopy type theory. We will leave aside this issue here and turn to the quest of giving meaning explanations for identity in intensional type theory in Chapter 5. Extending the meaning explanations to higher inductive types and universes is highly interesting, but beyond our analysis of the logical constants.

Following Martin-Löf's approach, we will first explain the meaning of the four kinds of judgments of type theory in Section 4.2.1. Afterwards, we will introduce types and terms into the meaning explanations by translation from HoTT into the simply theory of expressions in Section 4.2.2 and Section 4.2.3. The meaning explanations of the judgments give us a clear criterion for deciding if a type has been defined successfully, this criterion will be introduced in Section 4.2.4.

4.2.1 The meaning of the judgments

Of special status in the system of intuitionistic type theory are the so-called *canonical forms*, which can be either terms or types. According to Martin-Löf's conception of types, to know that A is a type is to know what counts as a canonical term of A and to know how to show that two canonical terms of the type A are equal[1]. The meaning of the four kinds of judgments introduced in Section 2.1 follows from this conception of a type.

1 It should be noted that this does not mean that we have to give rules that exhaustively produce all canonical forms — in some cases this is not possible, for example for functions.

Below in Section 4.2.2, we will introduce a translation of homotopy type theory into the simply typed lambda calculus, which we will call the *theory of expressions* in accordance with Dybjer (2012). Since the lambda calculus has a very simple syntax and primitive notion of computation, the translation should ease the project of justifying the rules of HoTT. In this translation, every term a of HoTT will be translated into a corresponding term $\mathtt{a}$ of the theory of expressions. The reference of a can then be retrieved by evaluating $\mathtt{a}$ to canonical form, say, $\mathtt{v}$. The canonical form relation is denoted $\mathtt{a} \Rightarrow \mathtt{v}$ and is vacuous for the time being, we will extend the relation as we consider all types of interest for us. Importantly, the canonical form of a term is unique, which goes well with the reading that the canonical form is the reference of the term (this understanding will be explained further in Chapter 6).

We can now formulate the meaning of the judgments in terms of the canonical form relation:

- The meaning of a judgement A *type* is that $\mathtt{A} \Rightarrow \mathtt{V}$.
- The meaning of a judgement $x : A$ is that $\mathtt{x} \Rightarrow \mathtt{v}$, $\mathtt{A} \Rightarrow \mathtt{V}$ and $v : V$.
- The meaning of a judgement $A \equiv B$ is that $\mathtt{A} \Rightarrow \mathtt{V}$ and $\mathtt{B} \Rightarrow \mathtt{V}$.[2]
- The meaning of a judgement $x \equiv y$ is that $\mathtt{x} \Rightarrow \mathtt{v}$ and $\mathtt{y} \Rightarrow \mathtt{v}$.

Intuitively, the meaning of A *type* is that A is a valid type, which is evident if A evaluates to a canonical form $\mathtt{V}$. In order to see if $x : A$ is meaningful, we have to evaluate both x and A and check if the canonical form of x is an inhabitant of the canonical form of A.

The judgmental equalities establish an equality of meaning. For example, if two terms $\mathsf{add}(\mathsf{s}(\mathsf{s}(0)), \mathsf{s}(\mathsf{s}(0)))$ and $\mathsf{double}(\mathsf{s}(\mathsf{s}(0)))$ both evaluate to the same term $\mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{s}(0))))$, we consider both terms to be equal in meaning and the judgmental equality to hold. Naturally, we want the judgmental equalities to be equivalence relations, which we can verify quickly:

- If we have $x : A$ such that $\mathtt{x} \Rightarrow \mathtt{v}$, $\mathtt{A} \Rightarrow \mathtt{V}$ and $v : V$, we obviously have $x \equiv x$.
- Assume $x \equiv y$, hence $\mathtt{x} \Rightarrow \mathtt{v}$ and $\mathtt{y} \Rightarrow \mathtt{v}$, i.e., $y \equiv x$.

2 Note that in the original presentation of the meaning explanations in Martin-Löf (1982), judgmental equality between types is defined extensionally:

> Two canonical types A and B are equal if a canonical object of type A is also a canonical object of type B and, moreover, equal canonical objects of type A are also equal canonical objects of type B. — Martin-Löf (1982):p. 163

We will follow Dybjer (2012) in having an intensional treatment of judgmental equality since it mirrors judgmental equality between terms. Furthermore, judgmental equality defined intensionally can be checked by simple type checking algorithms and hence is consistent with the idea that judgements should be *evident*.

- Assume $x \equiv y$ and $y \equiv z$. Hence, $\mathtt{x} \Rightarrow \mathtt{v}$ and $\mathtt{y} \Rightarrow \mathtt{v}$, and $\mathtt{y} \Rightarrow \mathtt{w}$ and $\mathtt{z} \Rightarrow \mathtt{w}$. Since canonical forms are unique, we infer that $\mathtt{v}$ and $\mathtt{w}$ must have been the same canonical forms, thus, $x \equiv z$.

The same explanation can be stated, *mutatis mutandis*, for judgmental equality between types. Granström (2008):p. 74 introduces the difference between *meaning determining inference rules* and *justified inference rules*, where the former rules are evident directly and the latter rules require further justification; for our definition of the meaning of the judgments, all properties of $\equiv$ follow directly so we do not need to introduce any difference in epistemological immediacy.

The above judgments were all made in the empty context, they are called *categorical judgements*. The meaning explanations of categorical judgments are extended to *hypothetical judgments*, i.e., judgements which are made in a context. To argue that hypothetical judgements are meaningful, we need to carry out an induction on the number of assumptions in the context. We will not repeat the argument here and refer to Martin-Löf (1982):p. 163–166 for an exposition of how to extend the meaning of the categorical judgements to hypothetical judgements. For us it will suffice to read a judgment such as

$$x : A \vdash B(x) \ type$$

as "in context $x : A$, $B(x)$ is a valid type".

4.2.2 The theory of expressions

We will translate the terms and types of HoTT into the simply typed lambda calculus as defined in Section 3.3, which we will call the *theory of expressions*. Both terms and types of HoTT will be represented as terms in the theory of expressions. We will write the interpreted terms in typewriter font, e.g., $\mathtt{a}$, to distinguish them from the terms and types of HoTT, e.g., a. For all type and term symbols, we introduce the following constants in the calculus, typed over a set of atomic types consisting only of type ι:

$$\mathbf{0} : \iota,\ \mathbf{1} : \iota,\ \star : \iota, \mathtt{R}_{\mathbf{1}} : \iota \to \iota \to \iota,$$

$$\mathtt{N} : \iota,\ \mathtt{s} : \iota \to \iota,\ \mathtt{0} : \iota,\ \mathtt{R}_{\mathtt{N}} : \iota \to \iota \to (\iota \to \iota \to \iota) \to \iota$$

$$\prod : \iota \to (\iota \to \iota) \to \iota,\ \lambda : (\iota \to \iota) \to \iota,\ \mathtt{App} : \iota \to \iota \to \iota$$

$$\sum : \iota \to (\iota \to \iota) \to \iota,\ \mathtt{pair} : \iota \to (\iota \to \iota) \to \iota,\ R_{\Sigma} : \iota \to (\iota \to \iota \to \iota) \to \iota$$

$$+ : \iota \to \iota \to \iota,\ \mathtt{inl} : \iota \to \iota,\ \mathtt{inr} : \iota \to \iota,\ \mathtt{R}_{+} : \iota \to (\iota \to \iota) \to (\iota \to \iota) \to \iota$$

Note that **0** and **1** denote the empty type and the unit type, respectively, while $\mathtt{0}$ denotes the 0 symbol of the natural numbers.

The type assignments can be understood as follows: A dependent function type $\prod_{x:A} P(x)$ takes an element of type A, here simply ι, a dependent type $P : A \to \mathcal{U}$, represented $\iota \to \iota$, and yields a type in the universe, again of type ι. Hence, the type of $\prod(\mathtt{A}, (\mathtt{x})\mathtt{P})$ is $\iota \to (\iota \to \iota) \to \iota$.

The dependent sum type $\sum_{x:A} P(x)$ takes as input a type A and a dependent type $P : A \to U$, and yields a dependent pair (a, p) with $a : A$ and $p : P(A)$. We will represent pairs also as basic entities in the universe and type them with ι. Hence, we type $\sum(\mathtt{A}, (\mathtt{x})\mathtt{P})$ with $\iota \to (\iota \to \iota) \to \iota$. The type of `pair` stems from the fact that we will represent dependent pairs, i.e., the second element of the pair is a function that takes as input the first element of the pair and hence, the type of `pair` is $\iota \to (\iota \to \iota) \to \iota$.

Recall that we can define function and product types as non-dependent cases of dependent function and dependent pair types, hence, the above theory allows us to interpret a significant part of homotopy type theory consisting of all type formers corresponding to the logical constants.

Now that we have introduced the types, we define the terms of the theory of expressions:

$$\begin{aligned} \mathtt{a} ::= \mathbf{0} \mid \mathbf{1} \mid \star \mid \mathtt{N} \mid \mathtt{0} \mid \mathtt{s}(\mathtt{a}) \mid \mathtt{R}_{\mathtt{N}}(\mathtt{a}, \mathtt{a}, \mathtt{a}) \mid \\ \prod \mathtt{x} : \mathtt{a}.\mathtt{a} \mid \lambda(\mathtt{a}) \mid \mathtt{App}(\mathtt{a}, \mathtt{a}) \mid \sum \mathtt{x} : \mathtt{a}.\mathtt{a} \mid \mathtt{pair}(\mathtt{a}, \mathtt{a}) \mid \mathtt{R}_{\Sigma}(\mathtt{a}, \mathtt{a}) \mid \\ \mathtt{a} + \mathtt{a} \mid \mathtt{inl}(\mathtt{a}) \mid \mathtt{inr}(\mathtt{a}) \mid \mathtt{R}_{+}(\mathtt{a}, \mathtt{a}, \mathtt{a}) \end{aligned}$$

Recall again that in the theory of expressions, abstraction is written $(\mathtt{x})\mathtt{a}$ and application is written $\mathtt{f}(\mathtt{a})$. In the definition of the grammar, we have written $\mathtt{f}(\mathtt{a}_1, ..., \mathtt{a}_\mathtt{n})$ instead of $\mathtt{f}(\mathtt{a}_1)...(\mathtt{a}_\mathtt{n})$. We have also used the more intuitive notations $\prod \mathtt{x} : \mathtt{A}.\mathtt{B}$ instead of $\prod(\mathtt{A}, (\mathtt{x})\mathtt{B})$, $\sum \mathtt{x} : \mathtt{A}.\mathtt{B}$ instead of $\sum(\mathtt{A}, (\mathtt{x})\mathtt{B})$, and $\mathtt{a} + \mathtt{a}$ instead of $+(\mathtt{a}, \mathtt{a})$.

In the following, we will use various kinds of letters as variables in the calculus to ease the legibility of the rules, including letters with subscripts such as $\mathtt{z}_0$.

The translation of HoTT into the theory of expressions is straightforward. We will only give a few examples:

$$\begin{aligned} 0 &\leftrightsquigarrow \mathtt{0} \\ \mathtt{s}(a) &\leftrightsquigarrow \mathtt{s}(\mathtt{a}) \\ \prod_{x:A} B(x) &\leftrightsquigarrow \prod \mathtt{x} : \mathtt{A}.\mathtt{B} \\ \lambda x : A.a &\leftrightsquigarrow \lambda((\mathtt{x})\mathtt{a}) \\ &... \end{aligned}$$

Note that expressions generated by the above grammar will always be finitary, e.g., there is no infinite term $...\mathtt{s}(\mathtt{s}(\mathtt{0}))$ representing infinity.

4.2.3 Canonical forms

In order to explicate the meaning explanations of the judgments as stated above in Section 4.2.1, we have to distinguish certain terms of the theory of expressions, which we will call the *canonical forms*. The canonical forms are exactly those terms that have been introduced with a type formation or term introduction rule. To be precise, the canonical forms are generated by the following grammar:

$$\begin{aligned} \mathtt{v} ::= &\ \mathbf{0} \mid \mathbf{1} \mid \star \mid \mathtt{N} \mid \mathtt{0} \mid \mathtt{s(a)} \mid \\ &\textstyle\prod \mathtt{x : a.a} \mid \lambda(\mathtt{a}) \mid \sum \mathtt{x : a.a} \mid \mathtt{pair(a,a)} \mid \\ &\mathtt{a + a} \mid \mathtt{inl(a)} \mid \mathtt{inr(a)} \end{aligned}$$

We will have $\mathtt{v}$, $\mathtt{w}$ and $\mathtt{V}$ denote canonical forms in the following, where v, w are terms and V is a type in HoTT. The term $\mathtt{a}$ denotes a closed, but not necessarily canonical term. Hence, only the outermost term constructor is relevant when determining if an expression is in canonical form, whereas the inner contents of a term do not need to be evaluated further. This evaluation strategy is known as "lazy" in functional programming. For example, $\lambda(\mathtt{(x)add(x,s(0))})$ is considered a canonical form, even though $\mathtt{(x)add(x,s(0))}$ is not in canonical form. Also, the terms $\mathtt{s(add(s(0),s(s(0))))}$ and $\mathtt{add(s(s(0)),s(s(0)))}$ are both not fully evaluated, but the first term is considered to be in canonical form since $\mathtt{s(a)}$ is in canonical form for any $\mathtt{a}$.

Non-canonical forms in the system reduce to canonical forms by employing the computation rules of the respective type formers. To represent this in our meaning explanations, we introduce the relation

$$\mathtt{a} \Rightarrow \mathtt{v}$$

between closed terms $\mathtt{a}$ and $\mathtt{v}$, which is read as "$\mathtt{a}$ has canonical form $\mathtt{v}$". We define the canonical form relation with inference rule notation:

$$\frac{}{\mathtt{v} \Rightarrow \mathtt{v}} \text{ (every canonical form has itself has canonical form)}$$

$$\frac{\mathtt{a} \Rightarrow \star \quad \mathtt{z} \Rightarrow \mathtt{v}}{\mathtt{R_1(a,z)} \Rightarrow \mathtt{v}}$$

$$\frac{\mathtt{n} \Rightarrow \mathtt{0} \quad \mathtt{z_0} \Rightarrow \mathtt{v}}{\mathtt{R_N(n,z_0,z_s)} \Rightarrow \mathtt{v}} \qquad \frac{\mathtt{m} \Rightarrow \mathtt{s(n)} \quad \mathtt{z_s(n,R_N(n,z_0,z_s))} \Rightarrow \mathtt{v}}{\mathtt{R_N(m,z_0,z_s)} \Rightarrow \mathtt{v}}$$

$$\frac{\mathtt{f} \Rightarrow \lambda(\mathtt{b}) \quad \mathtt{b(a)} \Rightarrow \mathtt{v}}{\mathtt{App(f,a)} \Rightarrow \mathtt{v}}$$

$$\frac{\mathtt{c} \Rightarrow \mathtt{pair(a,b)} \quad \mathtt{z(a,b)} \Rightarrow \mathtt{v}}{\mathtt{R_\Sigma(c,z)} \Rightarrow \mathtt{v}}$$

$$\frac{\mathtt{c} \Rightarrow \mathtt{inl(a)} \quad \mathtt{z_a(a)} \Rightarrow \mathtt{v}}{\mathtt{R_+(c,z_a,z_b)} \Rightarrow \mathtt{v}} \qquad \frac{\mathtt{c} \Rightarrow \mathtt{inl(b)} \quad \mathtt{z_b(b)} \Rightarrow \mathtt{v}}{\mathtt{R_+(c,z_a,z_b)} \Rightarrow \mathtt{v}}$$

Let us look at a few examples. Recall that the recursors are used to define functions in HoTT. For example, if we define

$$\mathsf{double}(x) :\equiv R_N(x, 0, \lambda n.\lambda y.\mathsf{s}(\mathsf{s}(y))),$$

we can evaluate double(s(0)) to canonical form by unfolding the definition of double and translating the resulting term into the theory of expressions:

$$\mathtt{R_N}(\mathtt{s}(0), 0, \lambda((\mathtt{n})\lambda((\mathtt{y})\mathtt{s}(\mathtt{s}(\mathtt{y})))))$$

This term can be evaluated by applying the first reduction rule of the natural numbers recursor, which requires us to first evaluate

$$\mathtt{R_N}(0, 0, \lambda((\mathtt{n})\lambda((\mathtt{y})\mathtt{s}(\mathtt{s}(\mathtt{y}))))),$$

which yields 0 if we apply the first reduction rule for $\mathtt{R_N}$. Plugging this result in the previous recursor, we retrieve the expected result:

$$\mathtt{R_N}(\mathtt{s}(0), 0, \lambda((\mathtt{n})\lambda((\mathtt{y})\mathtt{s}(\mathtt{s}(\mathtt{y}))))) \Rightarrow \mathtt{s}(\mathtt{s}(0))$$

All instances of the recursors will be eliminated by the above computation rules and all terms are reduced to canonical forms in the theory of expressions. By similar unfolding we can see that s(s(s(s(0)))) and s(add(s(0), s(s(0)))) are in canonical form, and add(s(s(0)), s(s(0))) computes to canonical form s(s(s(s(0)))).

4.2.4 The validity of the types

Now that we have examined the meaning of the judgements and we know when a judgement is made successfully, we can justify the inference rules of Martin-Löf's type theory. In Martin-Löf (1982), there is very little elaboration of why the inference rules are sensible:

> For each of the rules of inference, the reader is asked to try to make the conclusion evident to himself on the presupposition that he knows the premises. [...] there are also certain limits to what verbal explanations can do when it comes to justifying axioms and rules of inference. In the end, everybody must understand for himself.
>
> — Martin-Löf (1982)

In Martin-Löf (1984), there is more explanation of the inference rules. Common to all justifications is the idea that the inference rules should *preserve the meaning of the judgments*, i.e., it should remain evident that a judgment holds after new inference rules for a type are given. The meaning of the judgments relies on the fact that all terms are reducible to canonical forms, so newly introduced types needs to respect this property.

A type is introduced into the system by presenting the following set of rules[3]:

3 For some rules, additional uniqueness principles are stated to avoid distinguishing terms that are morally equal, such as f and $\lambda x.f(x)$. This is known as η-reduction, we will not be concerned with it in the following.

- A *formation rule* specifying how new types with the type symbol can be built.
- One or several *introduction rules* specifying how new inhabitants of the type can be introduced.
- One or several *elimination rules* stating how an inhabitant may be used.
- One or several *computation rules* specifying what happens when elimination rules are applied to inhabitants of the type.

We have used the computation rules already in the definition of the canonical forms above in Section 4.2.3. They encode the whole behaviour since they bind together introduction and elimination rules. If we specify how any term of a type inhabitant may be used, we have extended the system in a sound way. This is captured by the following criterion:

> **Successful definition of a type**:
> A set of inference rules for a type is successful iff:
>
> 1. For all A that can be constructed according to the formation rule of the type, we have a canonical type V such that `A` $\Rightarrow$ `V`.
> 2. For all $a : A$, we have a term $v : V$ such that `a` $\Rightarrow$ `v`.

In other words, we require that every type and every term that we have introduced in the system can be reduced to a canonical form in the theory of expressions. This property is also called *canonicity*.

A consequence of canonicity is that judgmental equality between terms, respectively types, is decidable. Hence, the above criterion ensures that newly introduced types do not impede the meaningfulness of all forms of judgments.

For example, let us consider the product type, which is a special case of the dependent sum type. If we have two inhabitants $(a, b) : A \times B$ and $(c, d) : A \times B$, then we should be able to recognize if both inhabitants are judgmentally equal. This is the case since in order to check $(a, b) \equiv (c, d)$, we can check $a \equiv c$ and $b \equiv d$ by reducing all terms to canonical forms and comparing them.

If we understand the canonical form as the *reference* of a term, the successful definition of types ensures means that all terms are meaningful, i.e., denote something. We will further inspect this understanding in the discussion in Chapter 6.

4.3 Breaking Canonicity and Propositional Canonicity

We have seen how intuitionistic logic can be modelled in the truncated version of HoTT in Section 4.1. This gives rise to what is called "intuitionistic constructivity" in HoTT Book (2013):Introduction. Another kind of constructivity, namely "algorithmic constructivity" is characterized by the meaning explanations and the insistence that all terms reduce to canonical forms. Under the traditional view, the antagonist of constructive reasoning is the law of excluded middle (LEM), and we will investigate in Chapter 6 how the law of excluded may be assumed in HoTT. However, constructivity in the algorithmic sense, i.e., the property that all proofs exhibit computational content, may be broken also in other ways than by applying the LEM. By postulating the existence of terms without giving a way to eliminate them, canonicity is broken, which we will investigate in Section 4.3.1. Afterwards, we will introduce an interesting formal result in Section 4.3.2: For some simple types such as the natural numbers, it is possible to obtain canonical forms of terms even if they contain axioms and hence do not reduce to canonical forms. This allows for recovering computational content from proofs which are not constructive in the algorithmic sense.

As a side note, proof theorists have been concerned with extracting computational content from classical proofs for quite some time. Kreisel has initiated the "unwinding program" to recover constructive content of seemingly non-constructive proofs (Kreisel, 1951; Kreisel, 1952). In a discussion of Kreisel's program, Feferman (1996) accredits only a limited success to Kreisel's program. However, it has some interesting applications in proof theory as exemplified by Kohlenbach and Oliva (2003) and Berger and Schwichtenberg (1995). The Dialectica interpretation of Gödel (1958) can also be used to extract proofs from non-constructive programs. Unfortunately, investigating the connections to homotopy canonicity lies behind the scope of this work.

4.3.1 Breaking canonicity

Intuitionistic type theory as presented in Section 4.2 enjoys canonicity since it is not possible to construct a term that does not reduce to a canonical form. However, we can break canonicity of the system by postulating the existence of a term without explicitly constructing it. For HoTT as presented in the HoTT Book (2013), univalence does not hold as a theorem. In order to use univalence in proofs, the existence of a term of the appropriate type needs to be postulated as follows:

$$\mathsf{ua} : (A \simeq B) \to \mathsf{Id}_{\mathcal{U}}(A, B)$$

This means that we purport to have a function that takes any witness of an equivalence to a witness for identity, even though we do not have constructed such a function. When a proof term contains ua, we cannot reduce it any further. This has practical repercussions, for example, Brunerie (2018) has proven that the 4-th homotopy group of the 3-sphere behaves like the integers modulo some natural number n, but since his proof employs the univalence axiom, it is not possible to reduce the n to a canonical numeral, in this case

2. Hence, using an axiom in a proof destroys the computational content of a proof.

This gives us two notions of non-constructivity at our hands: The existence of non-canonical forms, and proofs employing the law of excluded middle. The latter version will also break algorithmic constructivity since we need to assume LEM as an axiom, as we will see in Chapter 6.

4.3.2 Propositional canonicity

Since univalence does not hold as a theorem in HoTT, significant research activity is directed towards constructing a type theory where univalence gets a constructive characterization and does not need to be assumed as an axiom. There has been significant progress towards this goal with the development of cubical type theory.

Simultaneously, Vladimir Voevdosky has conjectured that a weakened form of canonicity also holds for HoTT where univalence is assumed as an axiom. The conjecture states that every inhabitant of the natural numbers $a : \mathbb{N}$, maybe containing the univalence axiom, is propositionally equal to a canonical form $n : \mathbb{N}$, i.e., there exists a proof $p : \mathsf{Id}_{\mathbb{N}}(a, n)$. The homotopy canonicity conjecture (called like that since propositional equality is interpreted as homotopical equivalence in HoTT) was proved recently by Christian Sattler[4]. It is an open question whether a constructive proof of the conjecture can be given, i.e., if there is a procedure that transforms any $a : \mathbb{N}$ into a canonical $n : \mathbb{N}$. Furthermore it is unclear if homotopy canonicity can be extended beyond the natural numbers to all set-like types.

Of interest for us is that homotopy canonicity shows that it is possible to retrieve computational content from proofs not evidently reducible to a canonical form. If the homotopy canonicity conjecture can be proven constructively, this retrieval can be carried out mechanically. If we understand $\mathtt{a} \Rightarrow \mathtt{v}$ as a having the reference v, then homotopy canonicity may have repercussions for our understanding of computational content and Martin-Löf's meaning explanations. We will further elaborate on this question in Chapter 6.

4.4 Proof-theoretic Semantics and the Meaning Explanations

Martin-Löf's approach to justifying intuitionistic type theory has been called the "syntactico-semantical approach to meaning theory" (Dybjer, 2012:p. 217). In contrast to the traditional view according to which syntax and semantics should be kept apart, intuitionistic type theory is not supposed to be without any content. Martin-Löf states in one of his lectures:

4 There exists no published proof yet, the result was presented at the conference HoTT-UF at the Center of Advanced Study in Oslo in June 2019.

> We will avoid keeping form and meaning (content) apart. Instead we will at the same time display certain forms of judgement and inference that are used in mathematical proofs and explain them semantically. Thus we make explicit what is usually implicitly taken for granted.
>
> — Martin-Löf (1984):p. 3

This is perpetuated in the portrayal of Martin-Löf's philosophy by Sundholm:

> [...] constructive type theory is an interpreted formal language. [...] the expressions used are real expressions that carry meaning. In a nutshell, the language is endowed with meaning by turning the proof-theoretic reductions into steps of meaning explanation. Just like the formulae of Frege's ideography, or of the language of Principia Mathematica, the type-theoretic formulae are actually intended to say something.
>
> — Sundholm (2012):p. xx

Similar convictions can be found in the project of *proof-theoretic semantics*, which has been developed in recent decades by Dag Prawitz and Peter Schroeder-Heister, among others. The main idea of proof-theoretic semantics is to specify the meaning of formal languages not by looking at an external system, but by investigating the formal system itself. It is closely related to the development of intuitionism, since both schools of thought understand the meaning of expressions by investigating the conditions under which they are *provable*, and not by investigating the conditions under which they are *true*. In order to better understand Martin-Löf's meaning explanations, we will look at some connections between proof-theoretic semantics and intuitionistic type in the following.

Proof theory and intuitionism are closely intertwined. Kreisel for example considered the central challenge of proof theory to elucidate the "mapping" between mental acts and derivations. Prawitz' *general proof theory* can be seen as combining an inferentialistic view on logic with an intuitionistic conception of mathematics. One reason for the confluence of inferentialism and intuitionism is that proof theory is biased towards intuitionistic logic since the main tool for proof-theorists, natural deduction, gives naturally rise to intuitionistic logic[5].

5 Schroeder-Heister (2018) invokes that classical logic in natural deduction violates the subformula property and has multiple logical constants in some inference rules. Additionally, it violates the counterpart of canonicity in natural deduction: The rules for classical logic do not ensure that any derivation be reduced to one which uses an introduction rule in the last step. Intuitionistic logic satisfies all these requirements.

4.4 PROOF-THEORETIC SEMANTICS AND THE MEANING EXPLANATIONS

The meaning explanations of Martin-Löf single out canonical forms as distinguished terms of the theory. If a new type is introduced, it has to be ensured that every term that can be constructed is reducible to a canonical form. This is closely related to the notion of *proof-theoretic validity* as proposed by Prawitz (1973); Prawitz (1974). Schroeder-Heister (2018) characterizes proof-theoretic validity as incorporating the following principles:

- The priority of closed canonical proofs.
- The reduction of closed non-canonical proofs to canonical ones.
- The substitutional view of open proofs.

The first two principles play a crucial role in Martin-Löf's meaning explanations. We have reduced the meaning of all non-canonical terms to that of canonical forms. The third point of determining the meaning of open proofs by substituting open variables by closed terms has not been worked out by Martin-Löf, but has been implemented by Dybjer (2012). Dybjer considers program testing procedures as meaning explanations for intensional type theory and specifies procedures that allow for injecting closed terms into open terms to check the validity of open terms.

According to the proof-theoretic view, inference rules are valid if they only allow for deriving constructively valid proofs from constructively valid proofs. The successfully defined types as defined in Section 4.2.4 are exactly those that ensure that any constructible term may be reduced to canonical form. Hence, valid proofs correspond to canonical forms in Martin-Löf's system. According to Schroeder-Heister, valid proof have the following special status:

> The definition of validity singles out those proof structures which are 'real' proofs on the basis of the given reduction procedures.
>
> — Schroeder-Heister (2018)

Crucially, proofs and objects are conflated in intuitionistic type theory. For example, a natural number has the same status as the proof of a mathematical proposition. The unifying view on objects and proofs means that not only derivations can be considered valid as in proof-theoretic semantics, but that we can also consider objects to be canonical. In the same way as a canonical proof represents the "most direct proof", the presentation of the natural number 4 as $\mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{s}(0))))$ is more "direct" than the presentation $\mathsf{add}(\mathsf{s}(\mathsf{s}(0)), \mathsf{s}(\mathsf{s}(0)))$. Intuitively, we would consider the answer to the question "How many automorphisms exist on the group $\mathbb{Z}_{12}^{+}$?" to be 4, and not $2+2$. Hence, Martin-Löf's meaning explanations give rise to distinguishing special mathematical objects as the "real" objects, and to determine the meaning of other expressions with respect to those distinguished objects.

4.4 PROOF-THEORETIC SEMANTICS AND THE MEANING EXPLANATIONS

In conclusion, there is significant overlap between Martin-Löf's meaning explanations and the ideas prevalent in proof-theoretic semantics. In contrast to most other work in proof-theoretic semantics, proofs are given an ontological character in type theory by treating them on equal footing with mathematical objects — this can be seen as the most resolute execution of *proof* theory, since it treats proofs as first-order mathematical objects.

Chapter 5

Constructive Identity

The original formulation of BHK, as portrayed for example in Troelstra and Dalen (1988), has no clause for identity. In Martin-Löf's intuitionistic type theory, every proposition is expressed by a type and consequently, identity is represented with a dedicated primitive type, the identity type. The elimination principle for identity types, commonly referred to as the *J-rule*, stipulates the existence of inhabitants of the identity type that are not in canonical form — which suggests that the identity type is set up unsuccessfully according to our exposition in Section 4.2.4. This behaviour was not intended by the creator of intuitionistic type theory, which led Martin-Löf to work on an extensional version of type theory which conflates propositional equality and judgmental equality (Martin-Löf, 1982). However, this conflation provokes that the judgments are not decidable anymore, such that Martin-Löf returned to working out intensional type theory afterwards. We will lay out how Martin-Löf's meaning explanations invalidate identity in both intensional and extensional type theory in Section 5.1. Martin-Löf was not able to answer the following question: How can the unusual structure of the identity type be explained?

The first insightful explanation of the J-rule has been carried out by regarding it as *path induction* in a topological space. The homotopy interpretation of type theory is the fundamental insight of the HoTT Book (2013) and allowed for a highly fruitful use of Martin-Löf's type theory. It seems auspicious to draw from this connection to give better meaning explanations for intensional type theory, we will sketch how this might be done in Section 5.2.

When defining identity in intensional type theory, Martin-Löf shows as a direct corollary that "the law of equality corresponding to Leibniz's principle of indiscernibility holds, namely that equal elements satisfy the same properties" (Martin-Löf, 1984:p. 61). It seems that Martin-Löf only set out to incorporate this very principle in his type theory when defining identity and formulating the J-rule. This raises the question if indiscernibility is enough to understand identity in intensional type theory and if the correspondence between identity and indiscernibility might give rise to an adaption of the meaning explanations which do validate identity. Ladyman and Presnell argue in their 2015 paper that indiscernibility does justify the "epistemological and methodological status" of the J-rule, but work out in their 2017 paper that the identity type cannot be derived from a type expressing indiscernibility — which is unsurprising since the predicativity of Martin-Löf's type theory impedes that the identity type can be defined without introducing primitive rules in the system. We will reconstruct the arguments of Ladyman and Presnell (2015) and Ladyman and Presnell (2017) in Section 5.3 and argue that indiscernibility is an unlikely candidate to patch Martin-Löf's meaning explanations since it cannot elucidate the behaviour of identity in intensional type theory.

We will close with an outlook on what a constructive justification of identity could look like, and a digression to the relevance of univalence for our project in Section 5.4.

One other project should be mentioned in this context: Walsh (2017) works on justifying the J-rule from a different perspective, namely by arguing that the behaviour of the J-rule can be understood from categorical grounds. He gives a new criterion of proof-theoretic harmony based on adjoints, a ubiquitous concept in category theory. The newly introduced notion of *categorical harmony* does indeed validate the J-rule. While this is a highly interesting explanation of identity in HoTT, we focus on the *constructive* validity of identity in in our work and will not go into more detail of the work of Walsh (2017).

Another recent work on the justification of the J-rule has been brought forward by Klev (2017) — however, his meaning explanations require that there are no non-canonical proofs of the identity type and hence, his approach is not compatible with univalence and homotopy type theory.

5.1 Identity in Martin-Löf's Meaning Explanations

The treatment of identity has changed throughout Per Martin-Löf's works: In Martin-Löf (1975), we can find the first exposition of intuitionistic type theory with an identity type representing propositional equality. Here we can also find the J-rule for the first time. Martin-Löf (1982) gives the first exposition of the meaning explanations of the judgments, but conflates propositional and judgmental equality. Martin-Löf (1984) presents a system without further committing to a specific set of rules and gives more explicate motivations for the rules. However, there is a mismatch between these motivations and the properties of identity in intuitionistic type theory. Consider the following two rules:

$$\frac{p : \mathsf{Id}_A(a, b)}{a \equiv b} \quad \text{(ER)}$$

$$\frac{p : \mathsf{Id}_A(a, b)}{p \equiv \mathsf{refl}} \qquad \text{(UIP)}$$

The rule ER conflates the two equalities of type theory since it enables us to conclude from the proof of a propositional equality a judgmental equality (the converse of ER is true in any version of Martin-Löf's type theory). The principle UIP states that the identity type yields no interesting structure since all identity proofs are equal to the trivial identity proof. Note that in this rule, the canonical element refl does not depend on a or b (as it is the case in Section 2.4), since a and b are judgmentally equal and there is only one unique proof of identity. A version of intuitionistic type theory validating both principles ER and UIP is commonly called *extensional*, whereas a type theory invalidating ER is called *intensional*. Homotopy type theory furthermore invalidates UIP since univalence requires the existence of non-trivial identity proofs[1].

In Section 5.1.1, we will retrace why Martin-Löf's meaning explanations, as presented in Section 4.2, are in conflict with intensional type theory since it invalidates UIP. In Section 5.1.2, we will argue that the meaning explanations also invalidate an extensional treatment of identity since ER is incompatible with the requirement that judgments should be evident.

It should be noted that the meaning explanations do not lead to trouble for simple types decidable equality. For example, propositional equality in the natural numbers is decidable, i.e., we can give an inhabitant of the following type:

$$\prod_{m,n:\mathbb{N}} \mathsf{Id}_{\mathbb{N}}(m, n) + (\mathsf{Id}_{\mathbb{N}}(m, n) \to \mathbf{0})$$

It is not a problem for the natural numbers that judgmental and propositional equality are conflated since we are still able to decide all judgments. It has been shown by Hedberg (1998) that UIP holds for all types in which equality is decidable. Hence, all such simple types do not pose a problem for Martin-Löf's meaning explanations — neither to extensional type theory (since judgments continue to be decidable), nor to intensional type theory (since there are no non-trivial identity proofs).

5.1.1 Intensional type theory and the meaning explanations

In order to extend the meaning explanations given in Section 4.2 with identity, we could try to introduce the following constants in our theory of expressions:

$$\mathtt{Id} : \iota \to \iota \to \iota, \ \mathtt{refl} : \iota \to \iota, \ \mathtt{R}_{=} : \iota \to \iota \to \iota$$

Note that we make `refl` dependent on a specific element since ER is not validated. The grammar of the theory of expressions could then be extended

1 See, e.g., HoTT Book (2013):Example 3.1.9

as follows:

$$\mathtt{a} ::= \ldots\ \mathtt{Id}(\mathtt{a},\mathtt{a}) \mid \mathtt{refl}(\mathtt{a}) \mid \mathtt{R}_{=}(\mathtt{a},\mathtt{a})$$

The translation of HoTT into the theory of expressions is straightforward.
As canonical forms we consider the identity type and the trivial-self identification:

$$\mathtt{a} ::= \ldots\ \mathtt{Id}(\mathtt{a},\mathtt{a}) \mid \mathtt{refl}(\mathtt{a})$$

The computation rule for identity eliminates the recursor of identity, thereby we extend the canonical form relation with the following rule:

$$\frac{\mathtt{c} \Rightarrow \mathtt{r} \quad \mathtt{t} \Rightarrow \mathtt{v}}{\mathtt{R}_{=}(\mathtt{c},\mathtt{t}) \Rightarrow \mathtt{v}}$$

Univalence implies that we have a type A with $a, b : A$ such that $p : \mathsf{Id}_A(a,b)$, but neither $\mathtt{p} \Rightarrow \mathtt{refl}(\mathtt{a})$ nor $\mathtt{p} \Rightarrow \mathtt{refl}(\mathtt{b})$. Hence, there is no canonical form $v : \mathsf{Id}_A(x,y)$ such that $\mathtt{p} \Rightarrow \mathtt{v}$. This is in conflict with our criterion for the successful definition of a type proposed in Section 4.2.4.

A possible explanation for why canonicity fails for identity types is that we have set up identity in the wrong way. Per Martin-Löf seemed to have shared this view for a few years, which is why he pursued a different version of his type theory in Martin-Löf (1982).

5.1.2 Extensional type theory and the meaning explanations

In extensional type theory, UIP is validated. Hence, there are no non-canonical identity proofs and the criterion for the successful definition of a type is satisfied. Martin-Löf's meaning explanations are commonly considered to validate extensional type theory[2] and this is true if we consider the informal explanations of the set of rules in Martin-Löf (1984) as the meaning explanations. However, extensional type theory invalidates the meaningfulness of the judgments: The judgments are not decidable anymore. We cannot decide judgmental equality with a mechanic procedure, and type-checking is also not decidable anymore. This is in stark contrast with the intuitionistic dogma:

> We recognize a proof if we see one.
>
> — Kreisel (1966):p. 202

Martin-Löf subscribed to the same paradigm, in Martin-Löf (1987) we can find the conviction that judgments should be *evident*. The best characterization of evidence is that we have a terminating mechanical procedure deciding judgments, but such a procedure does not exist for extensional type theory.

2 We can find this conviction, for example, in a talk of Peter Dybjer: https://www.youtube.com/watch?v=dJF-iW5Gav8.

We conclude that extensional type theory is also incompatible with the meaning explanations for intuitionistic type theory. Even though extensional type theory seems at first sight more "well-behaved" than intensional type theory, the lack of distinction between judgmental and propositional equality is not only impractical, but also has conceptual disadvantages over intensional type theory. The success of intensional type theory in the univalent foundations program should motivate us to pursue intensional type theory and justify its behaviour.

5.2 Homotopical Interpretation of Identity

Identity in intensional type theory seemed ill-behaved for a long time, implementations of the type theory in theorem provers often got rid of its unusual behaviour by assuming UIP. Hofmann and Streicher (1998) formulated a concise countermodel for UIP by showing that the identity type in intensional type theory exhibits the structure of a groupoid. This prepared the ground for the insight that the internal structure of types in intensional type theories can be modelled by ∞-groupoids, which in turn are a model of homotopy types. The connection was drawn independently by Awodey and Warren (2009) and Voevodsky (2006) and led to the formulation of many important new concepts such as univalence. Since the homotopy interpretation was so crucial in making use of intensional type theory, we will investigate in this section what lessons we can learn from this connection in order to justify the constructive validity of identity. We will briefly recap the homotopy interpretation in Section 5.2.1 and then discuss how this is relevant to our project of adapting Martin-Löf's meaning explanations in Section 5.2.2.

5.2.1 The homotopy interpretation

The fundamental insight developed in the HoTT Book (2013) is the profound correspondence between intuitionistic type theory and homotopy theory: Types can be regarded as abstract topological spaces, inhabitants of the type as points in the space and identifications between terms as paths between points.

Consider Figure 5.1: The type A is regarded as a space. Inhabitants $a : A$ and $x : A$ are points in that space (to be precise, functions from the respective single point into the space). The identity proofs $\mathsf{refl}_a : \mathsf{Id}_A(a, a)$ and $p : \mathsf{Id}_A(a, x)$ are interpreted as paths r and p in the space A, i.e., as functions from the unit interval $I = [0, 1]$ into the space with adequate endpoints:

$$r : I \to A \text{ with } r(t) = a \text{ for all } t \in I$$
$$p : I \to A \text{ with } p(0) = a,\ p(1) = x$$

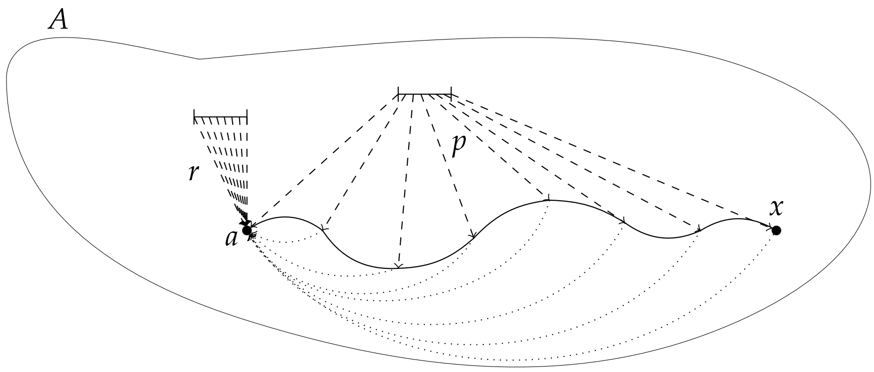

Figure 5.1: The homotopy interpretation of type theory

We can define a homotopy h, i.e., a continuous transformation, between r and p by $h(t, x) = p(tx)$. The transformation is depicted in Figure 5.1 with the dotted points: Every point on p can be mapped into the singe point a. Homotopies can be seen as the notion of identity in homotopy theory, every statement can only be made *up to homotopy*. This justifies why the analogue of the J-rule is valid in the homotopical picture: We want to prove that the property C holds for any point x and path p starting from point a. We have proved that C holds for the constant path r. Since r and p are homotopical, we can retract x along p to a. Hence, the property C also holds for p.

The name path *induction* stems from the fact that, informally speaking, we only need to prove that C holds for the base case of the constant path r. Since p is identical to r, the induction step is given to us for free.

Regular path induction, which is equivalent to based path induction, can be interpreted by a similar argument, where we vary both ends a and x of the path.

5.2.2 Meaning explanations based on the homotopy interpretation

The homotopy interpretation gives nice geometrical intuitions for the peculiarities of intuitionistic type theory and allows to reason about homotopical spaces in an axiomatic way. Can this be used to provide meaning explanations for intensional type theory?

If we interpret meaning explanations loosely in the sense of giving some meaning to the rules of type theory, the homotopy interpretation is adequate. Tsementzis (2019) abandons Martin-Löf's meaning explanations all together and goes on to develop a new kind of meaning explanations based on spatial notions such as "shape", "path" and "point". This can be seen as the attempt at articulating the metaphysical subject matter of type theory, similar to collections as the metaphysical underpinning of set theory. However, this is of little use for our project of formulating a constructive account of identity. The meaning explanations of Martin-Löf are concerned with explaining the computational content of constructive logic, where the puzzling point is that the principle that all proofs should reduce to some canonical form is not valid. Understanding why this is sensible and possibly adapting the definition of constructive validity are the main challenges we are facing. Discarding Martin-Löf's original meaning explanations seems like throwing out the baby with the bathwater since we want to retain his perspective on computational content — which does not mean that we should not take inspirations from the homotopical interpretation of type theory, but put bluntly, spaces and points are not the subject matter of logic.

5.3 Indiscernibility and Identity

One natural way to think about identity has been put forward by Leibniz and states that identity can be defined in terms of indiscernibility: If it is not possible to discern two entities by any property, the two entities must in fact be identical. When giving meaning explanations for type theory, it is natural to investigate if indiscernibility can illuminate the behaviour of propositional equality in HoTT. Ladyman and Presnell (2015); Ladyman and Presnell (2017) deal with this very question in the course of evaluating if HoTT can serve as foundation for mathematics. According to Ladyman and Presnell, a foundation should not only give methodological guidance, but is required to answer more questions regarding the ontology and epistemology of mathematics. In particular, all concepts in such a foundation have to be *justified* independently of mathematics, i.e., the concepts of the foundation must be derived from first principles. In particular, Ladyman and Presnell disregard path induction as a justification for the J-rule since it presumes notions from algebraic topology. The idea of a pre-mathematical justification stems from Linnebo and Pettigrew (2011), who introduced this criterion in the debate on whether category theory can serve as a foundation for mathematics.

The requirements that Ladyman and Presnell set for a pre-mathematical justification are more restrictive than what we could use as meaning explanations for identity in intensional type theory. In our presentation of the meaning explanations in Section 4.2, we have introduced another (albeit very basic) formal system to pin down the semantics of homotopy type theory and, *prima facie*, the homotopical interpretation is not ruled out as a meaning explanation for identity in intensional type theory.

So while the premises and motivation of Ladyman and Presnell are different from our project of justifying the constructive validity of identity in intensional type theory, their results might still provide insight in the features of identity and is highly relevant for our project. In particular, a successful justification should also explain why there are non-canonical identity proofs.

In the following, we will quickly introduce what has been called *Leibniz' law* in Section 5.3.1 and then reconstruct the argument of Ladyman and Presnell (2015) in Section 5.3.2. We will conclude in Section 5.3.3 that, in fact, the argument is not convincing and the authors do not live up to their aspirations of providing a pre-mathematical justification. Relevant to our project, we will quickly recap Ladyman and Presnell (2017) and concur with them that indiscernibility only gives limited insight into identity in intensional type theory.

5.3.1 Leibniz' law

Going back to Leibniz, identity is traditionally defined by stating that two things are identical if and only if the share all their properties.[3] His principle gives a sufficient criterion for identity of two objects by stating that if they are indiscernible, they must be equal. The *principle of the identity of indiscernibles* can be phrased like this:

$$\text{If } P(x) \leftrightarrow P(y) \text{ for all properties } P\text{, then } x = y. \qquad (\mathrm{PII}_{\rightarrow})$$

In other words, if we cannot distinguish two elements, we should consider them identical. The converse of this principle states that if two objects are equal, they cannot be distinguished by any property, sometimes called the *principle of the indiscernibility of identicals*:

$$\text{If } x = y\text{, then } P(x) \leftrightarrow P(y) \text{ for all properties } P. \qquad (\mathrm{PII}_{\leftarrow})$$

Note that two judgmentally equal types or terms are interchangeable in all contexts and hence, $\mathrm{PII}_{\leftarrow}$ holds without any constraints for judgmental equality. The converse, $\mathrm{PII}_{\rightarrow}$, does not hold for judgmental equality since one can not *prove* a judgmental equality by investigating the properties of types or terms. Judgmental equality should be thought of as the equality evident from symbol manipulation and has a distinctively intensional character. Hence, it is not meant for investigating equality between mathematical objects and cannot be explained with indiscernibility.

Propositional equality on the other hand characterizes non-trivial identifications between objects and is obtained by an explicit proof object. We will now investigate if it can be characterized satisfactorily by Leibniz' law, for which we will first investigate if the J-rule corresponds to $\mathrm{PII}_{\leftarrow}$. In combination with univalence, propositional equality can be thought of as a highly extensional equality that identifies structurally equal objects, which suggests that it incorporates $\mathrm{PII}_{\rightarrow}$, which we will discuss in Section 5.4.

3 Strictly speaking, only the principle of the identity of indiscernibles has been formulated by Leibniz. Since we are not concerned with historic cherry picking here, we will also attribute the converse of this principle to Leibniz.

5.3.2 Justifying the J-rule

In order to pre-mathematically justify the J-rule in HoTT, Ladyman and Presnell (2015) attempt to derive the J-rule from two primitive concepts which they consider comprehensible from intuitive grounds. We will introduce the two principles in the following.

We will consider an arbitrary type A and $a : A$ in the following and derive based path induction for a. We introduce the abbreviation $E(a)$ to refer to all inhabitants of A that are equal to a:

$$E(a) :\equiv \sum_{x:A} \mathsf{Id}_A(a, x)$$

Recall that the elements of a sigma type $\sum_{x:A} B(x)$ are tuples (x, p) where $x : A$ and $p : B(x)$, i.e., when viewing the type as a proposition, x is the witness of the existential statement and p the proof that x is indeed a witness. Inhabitants of $E(a)$ are tuples (b, p) where $b : A$ and $p : \mathsf{Id}_A(a, b)$.

The first principle relevant for the argument is the *uniqueness principle for identity types* (not to be confused with the *uniqueness of identity proofs* discussed above in Section 5.1). In general, uniqueness principles can be seen as sanity checks to ensure that a type has been defined appropriately: A uniqueness principle asserts that the only inhabitants of a type are those which can be constructed from the term constructors. For identity, the uniqueness principle reads as follows:

$$\prod_{(b,p):E(a)} \mathsf{Id}_{E(a)}((a, \mathsf{refl}_a), (b, p)) \tag{UPI}$$

Intuitively, UPI states that all elements equal to a should be equal to a, and the proof of that fact should be equal to the trivial self-identification refl_A. Crucially, this does not mean that there is only one inhabitant of $E(a)$ as Ladyman and Presnell point out — but only one inhabitant "up to identity" (Ladyman and Presnell, 2015:p. 403). Hence, this seemingly innocent principle does have a very subtle meaning, which we will further examine in the discussion in Section 5.3.3.

The second ingredient for the derivation of based path induction is what Ladyman and Presnell call *substitution salva veritate*. For this, let $Q : B \to \mathcal{U}$ be an arbitrary predicate over the type B. The following type expresses that if two elements s, t are proven to be equal and there is an inhabitant of $Q(t)$, then we also have an inhabitant of $Q(s)$:

$$\prod_{s:B} \prod_{t:B} \mathsf{Id}_B(s, t) \to Q(t) \to Q(s) \tag{SSV}$$

Read as a proposition, this is just $\mathrm{PII}_{\leftarrow}$ from above, the indiscernibility of identicals. Ladyman and Presnell claim that since indiscernibility is an adequate account of identity, we can introduce a term of the above type. We will an inhabitant of the above type in SSV $\mathsf{ssv}_{B,Q}$, indexed by B and the predicate Q.

If we consider $E(a)$ for type B and plug in (b, p) and (a, refl_a) into ssv, we have a function

$$\mathsf{ssv}_{E(a),Q}((b,p),(a,\mathsf{refl}_a)) : \mathsf{Id}_{E(a)}((b,p),(a,\mathsf{refl}_a)) \to Q(a,\mathsf{refl}_a) \to Q(b,p)$$

Given that UPI also holds, we can construct the following function:

$$Q(a, \mathsf{refl}_a) \to Q(b, p)$$

Hence, for any predicate Q and a witness that Q holds for a and the trivial self-identification refl_a, we can produce a proof of Q for any other inhabitant of the identity type. This corresponds to based path induction.

5.3.3 Discussion

We want to argue that the argument that "based path induction can be justified on the basis of pre-mathematical principles", as claimed and conducted by Ladyman and Presnell (2015), is not valid.

Crucial defect of the approach of Ladyman and Presnell (2015) is that they cannot motivate the uniqueness principle of identity types. While the principle of indiscernibility of identicals, introduced as SSV above, can very well be argued to only rely on pre-mathematical intuitions of what identity is, the uniqueness principle of identity types UPI presumes the intricate structure exhibited by identity in HoTT. As Ladyman and Presnell (2015) correctly acknowledge, UPI does *not* state that there is only one inhabitant $\sum_{x:A} \mathsf{Id}_A(a, x)$, namely (a, refl_A), but that this is the only inhabitant *up to identity*. This unusual notion is borrowed from the homotopical formulation "up to homotopy", and is highly surprising without knowing about the homotopy interpretation.

Indiscernibility of identicals is a corollary of the J-rule as shown in Section 2.4, but the J-rule says something different than $\text{PII}_{\leftarrow}$. In particular, indiscernibility does not give insight in why there are several inhabitants of the identity type and why those in turn can be identified. An obstacle seems to be that Ladyman and Presnell look at the elimination rule of identity in isolation, but we would need to make sense of all characteristics of identity to justify the J-rule specifically.

In Ladyman and Presnell (2017), the authors try to reduce identity to indiscernibility in a more direct way. While the goal pursued by Ladyman and Presnell (2015) only required the *motivation* of the J-rule from pre-mathematical grounds, the 2017 paper demands to directly recover the properties of the identity type from indiscernibility, which is defined as follows:

$$\mathsf{Indis}_A(a, b) :\equiv \prod_{P:A \to \mathcal{U}} P(a) \leftrightarrow P(b)$$

The authors are not able to derive the properties of the Id type from the definition above. This does not come as a surprise: In predicative type theories such as Martin-Löf's type theory, quantification over all predicates is not strong enough to allow for the definition of the identity type, which is why Martin-Löf had to introduce it as a primitive type with dedicated introduction, elimination and computation rules.

It should be noted that it is possible to define identity via Leibniz' law in another class of type theories, namely those with *impredicative* universes, i.e., where the universe of all types contains itself. The calculus of constructions (CoC) (Coquand and Huet, 1986) is such an impredicative type theory, a variant of the CoC forms the basis of the popular theorem prover Coq. In the CoC, identity does not need to be introduced as a primitive type, but can be defined similarly to the definition of Indis above.

We conclude that even though indiscernibility explains some aspects of identity in intensional type theory, it gives little insight in why there are several, non-canonical inhabitants of the identity type. Hence, indiscernibility underdetermines the characteristics of the identity type in HoTT and can only be part of the motivation for it. This is also recognized by Ladyman and Presnell (2017):

> Hence we conclude that, as is standardly supposed in work on HoTT, "identity" types really do express identity. So, unavoidably, this potential new foundation for mathematics has at its heart a radical new way of thinking about identity, which throws up new questions to be studied and, potentially, unexpected new applications to the many philosophical problems involving identity.
>
> — Ladyman and Presnell (2017):p. 242

5.4 Outlook

We have seen that neither Martin-Löf's meaning explanations, nor the characterization of identity as indiscernibility gives a satisfying answer to why there are non-canonical inhabitants of identity types. The homotopical interpretation does give an answer to that question by drawing from geometric intuitions, but while this is very fruitful for formalizing topology, it cannot justify the constructive character of identity in intensional type theory.

It seems that the criterion for a successful definition of a type as proposed in Section 4.2.4 needs refinement to incorporate identity types. We have seen in Section 5.3.2 that the uniqueness principle for identity types plays a crucial role for the motivation of the J-rule. UPI essentially ensures the uniqueness of identity proofs on the level of *propositional* equality. An ad-hoc adaption of the meaning explanations would be to extend the canonical form relation to propositional equality. This would ensure that all terms have a canonical form, but would destroy the decidability of the judgments, akin to the problems that extensional type theory face. Nevertheless, it seems that we have to weaken the insistence on canonical proofs to justify the structure of the identity type.

Ladyman argues that objects are only treated "up to identity" — this notion is exactly what we need to explain, since it only makes sense in the homotopical interpretation. Merely changing the terminology from paths to "computational deformations" seems to give little insight, but might be used as a basis for a constructive interpretation of identity.

The complex behaviour of identity in constructive type theory is not just legerdemain: Proof-irrelevant systems are only able to characterize decidable equality of simple sets, whereas intensional type theory allows for representing mathematical objects in a more fine-grained way. It seems inevitable that equality evinces an intricate structure if one asks:

In which way are things equal?

Univalence as identity of indiscernibles?

We will close with a short digression on the relevance of univalence for our project. Roughly speaking, univalence can be regarded as the principle that "isomorphic objects can be identified". This sounds a lot like the principle of the identity of indiscernibles stated in Section 5.3.1 — isomorphic objects cannot be discerned by any property, so we consider them the be equal. Unfortunately, we do not have enough time to further investigate univalence in this work, but it is worthwhile to assess the role of univalence for justifying the structure of identity in intensional type theory.

At first glance, the J-rule and univalence seem to be converse principles: The J-rule is the elimination rule of propositional equality, while univalence allows for establishing a propositional equality after observing that two things are equivalent. However, both principles work on different levels: Univalence is a property of a universe, hence it allows for identifying *types*. The J-rule in contrast is applied to properties and allows for using an identification between *terms*.

A constructive explanation of identity in HoTT does not necessarily have to incorporate univalence, since it is the structure induced by the identity type that we need to explain. However, the intense research effort on giving a constructive account to univalence might yield insight for our project, more investigation in the conceptual status of cubical type theories seems advisable for that purpose.

Chapter 6

Liberal Constructivism

Meaningful distinctions need to be preserved.

— Bishop (1973)

The distinction between constructive and non-constructive proofs is often not made clear, in fact, many mathematicians are unaware of when they employ proof steps that destroy the computational content of a proof. Brouwer and later Bishop have prominently called attention to this issue. In Bishop's view, constructive proofs are not only useful, but have a different status than non-constructive proofs since they directly give rise to mathematical objects satisfying the theorem under question. In Chapter 4, we have seen how Martin-Löf's meaning explanations attribute special status to canonical forms since they characterize the computational content of proofs. When developing a theory of meaning for mathematics, we want it to reflect the distinction between constructively valid and only classically valid propositions.

We also want more careful differentiation in another point: The fact that a statement cannot be proven constructively does not imply that the statement is invalid. Committed intuitionists in the tradition of Brouwer equate validity with constructability. Even though the focus on mental constructions gives a clear view on the epistemology of mathematics, it has a serious defect: Intuitionists have difficulty to explain why mathematicians converge in their ideas. Since the existence of a mind-independent mathematical universe is rejected, mathematical truths are not determinate. It seems difficult to explain why mathematics is so successful in creating an accepted body of knowledge if any mathematicians entertains her own mental constructions. We want to abandon the subjective view of intuitionism on mathematics in order to formulate a satisfying view on the semantics of mathematics. The assumption of a determinate universe is necessary if we want to maintain that the law of excluded middle (LEM) may be used coherently.

Already Bishop diverted from the intuitionistic conviction that the assumption of a mind-independent mathematical universe should be dismissed as "metaphysical". The view that a mathematical universe is explored by the mathematician is not only consistent, but necessary for his understanding of the nature of mathematics. As Goodman (1983):p. 65 puts it: "Bishop's insistence on objectivity ought to force him to accept classical logic", a view that is also supported by Beeson (1980): "the ontological viewpoint of Bishop is that mathematical objects (as well as mathematics in general) are *objective*". Bishop recognizes in the preface to his seminal *Foundations of constructive analysis* that the idealistic component of mathematics (namely the use of LEM) allows for simplifications and "opens possibilities which would otherwise be closed" (Bishop, 2012:p. viii). Bishop's work criticizes that classical mathematics neglects the importance of constructive reasoning, but it does not maintain that classically pursued mathematics is incorrect or incoherent.

It seems that nowadays, a significant part of researchers working in constructive mathematics share Bishop's lenient perspective on classical reasoning. The term *liberal constructivism* has been adopted by Billinge (2003) for this conviction, we will adopt this terminology. For example, Goodman has formulated the view of liberal constructivism as follows:

> [...] we must recognize that the theorems we prove are true about a determinate structure which we do not dream, but which is actual and independent of which of us is studying it.
>
> — Goodman (1983):p. 65

To the knowledge of the author, there has been no concise attempt at formulating a philosophy of mathematics that combines Goodman's stance with something that explains the special importance of constructive content of a proof. When investigating the writings of Bishop, Billinge (2003) concludes that Bishop did not bring forward a concise philosophy of mathematics. We want to provide an initial attempt at formulating such a philosophy of mathematics. In particular, we have to reconcile the different views on the semantics of mathematics, as intuitionists have a different view on the meaning of mathematical expressions than classically minded mathematicians. Is a proposition just a truth-apt sentence, or is it the set of its proofs? What is the reference of a theorem that has not been proven constructively? What do the logical constants mean? If we take as basic the view that has been indicated by Martin-Löf and is worked out in proof-theoretic semantics, then the meaning of the inference rules is determined solely by their use and not by referring to something outside of the formal system. However, applying the law of excluded middle presumes the existence of some determinate mathematical universe and requires to attribute the meaning of a mathematical statement to something outside of the formal system.

How can these views be reconciled? The different theories of meaning maintained by intuitionists and classical mathematicians seem incompatible at first glance, and it has been argued by Davies (2005) that constructive and classical mathematics are simply different frameworks with different subject matters. If the Univalent Foundations Program wants to succeed in providing a basis for all of mathematics, it will need to provide a unified semantical theory. An application of the Carnapian principle of tolerance is not acceptable if we want to create unified foundations for a fixed subject matter, in this case, mathematics. Hence there is a need for reconciliation between constructive and classical accounts to the meaning of mathematical expressions. A challenge will be that "any adequate *philosophical* defence of constructivism will also be an argument for the illegitimacy of classical mathematics" (Billinge, 2003:p. 150). We have to make the balance act of embracing constructive reasoning while maintaining that the LEM is a valid principle of reasoning.

In the HoTT Book (2013), it has been worked out how the law of excluded middle can be assumed consistently in intuitionistic type theory. We will trace this account in Section 6.1, where we will see why the LEM fails constructively and how assuming the LEM destroys the computational content of proofs. Afterwards, we will sketch a theory of meaning for HoTT that incorporates theorems proved by using the LEM in Section 6.2. We will discuss some open problems in Section 6.3 and conclude in Section 6.4, where we will argue that HoTT is very apt formal framework to realize the amalgamation of constructive and classical mathematics.

6.1 Law of Excluded Middle

As we saw in Chapter 3, the intuitionistic conception of logic did not result from devising a logic from first principles, but rather, intuitionistic logic has been developed to capture the canon of proofs that are acceptable from an intuitionistic point of view. Accordingly, intuitionists do not *a priori* regard the law of excluded middle as invalid, it is just that employing the LEM breaks the constructive content of proofs. We want to trace this incident in the context of HoTT in Section 6.1.1 and then investigate in Section 6.1.2 which commitments are carried by postulating the LEM in HoTT.

6.1.1 The constructive invalidity of LEM

The law of excluded middle states that each statement is either true or false. A formulation of LEM under a naïve propositions-as-types interpretation might be put forward as follows:

$$\mathsf{LEM}_\infty :\equiv \prod_{A:\mathcal{U}} A + (A \to \mathbf{0})$$

This formulation of LEM is inconsistent with univalence, see for HoTT Book (2013):Theorem 3.2.2. Actually, this is also not the formulation of LEM that we want: We want to claim that any proposition is either true or false, but not that any type is inhabited or empty. Otherwise, proving that A

cannot be inhabited and then applying LEM_∞ would yield a specific element of $A \to \mathbf{0}$ — but since there may be several inhabitants of that type, there cannot be a canonical way to obtain such an element.

As we have seen in Section 2.6, we can recover the classical understanding of propositions in HoTT by means of propositional truncation, yielding for any type A a truncated type $\|A\|$ that conflates all inhabitants to a unique inhabitant. Hence, any type can be transformed to a mere proposition, and all mere propositions are equal to either the empty or the unit type, i.e., mere propositions are truth-apt. With propositional truncation, we can give an adequate formulation of the LEM:

$$\mathsf{LEM} :\equiv \prod_{A:\mathcal{U}} \|A\| + (\|A\| \to \mathbf{0})$$

Type LEM expresses the proposition that any type, when regarded as a mere proposition without any computational content, can be proved or disproved. Under more considerate formulation, an inhabitant of LEM is an *effective procedure that produces for any type either an inhabitant of its truncated version or shows that, assuming there was an inhabitant, falsity can be derived*. Obviously, we cannot in general construct such a procedure for a formal system with the proof-theoretic strength of intuitionistic type theory. It is not a surprise the LEM is not a theorem in HoTT — instead, the constructive reading of the universal quantifier makes it necessary that the LEM is false.

It should be noted that for some types, the law of excluded middle does hold as a theorem, namely, for all type where there is an effective procedure that can decide a certain property. For example, equality between natural numbers is decidable (which we laid out in Section 5.1). Therefore, we do have an inhabitant of the following type in HoTT:

$$\prod_{m,n:\mathbb{N}} \mathsf{Id}_{\mathbb{N}}(m,n) + (\mathsf{Id}_{\mathbb{N}}(m,n) \to \mathbf{0})$$

This fact has already been pointed out by Tait (1994):p. 47.

6.1.2 Assuming the LEM

We have seen that the understanding of the LEM as an effective procedure deciding any proposition evinces that it cannot hold constructively. In this section we will see that it is nevertheless consistent to postulate the existence of such a procedure. If we assume the existence of an independent mathematical universe in which all propositions do have a determinate truth value, such a postulation is not harmful or incoherent. Hence, proving something classically in HoTT amounts to saying that *assuming that we would have an effective procedure deciding all propositions, we can derive the truth or falsity of a particular proposition*. This assumption is done by postulating that there is an inhabitant of LEM:

$$\mathsf{lem} : \mathsf{LEM}$$

Proof terms that contain lem will lack computational content, but since we are interested in propositions as truth-apt sentences, we cannot expect more than retrieving a truth value from a proof that has not been carried out constructively. Note that type-checking remains decidable even if the existence of lem is postulated. This means that we still adhere to the principle that we should recognize a proof when we see one.

Let us illustrate the application of the law of excluded middle in an example. The real numbers are traditionally subtle to introduce in constructive system, the HoTT Book (2013) presents two ways of constructing the real numbers by Cauchy sequences and Dedekind cuts. When the LEM is assumed, both formulations are equivalent in HoTT, so for the purpose of the example it does not matter which formulation of the real numbers we use, we refer to HoTT Book (2013):Ch. 11 for an introduction to the reals in HoTT.

Consider the intermediate value theorem, which states that for any function that is continuous on the unit interval and that changes its sign between 0 and 1, there exists a root. In HoTT, we can express the theorem as follows:

$$\mathsf{IV} :\equiv \prod_{f:[0,1]\to\mathbb{R}} [f(0) < 0 \wedge f(1) > 0) \to \sum_{x:[0,1]} f(x) = 0]$$

The canonical proof of the theorem proceeds by bisecting the interval step by step and hence approximating the root:

$$x_0 = \frac{1}{2}$$
$$x_n = \begin{cases} \frac{x_{n-1}}{2}, & \text{if } f(x_{n-1}) > 0 \\ \frac{1+x_{n-1}}{2}, & \text{if } f(x_{n-1}) < 0 \\ x_{n-1}, & \text{if } f(x_{n-1}) = 0 \end{cases}$$

The limit of the sequence $x_1, ..., x_n$ will the yield a value x for which $f(x) = 0$. Constructively, the proof does not go through since it is not possible to give an algorithm that decides if a real number is larger, smaller or equal to 0. Since real numbers are given as only potentially infinitesimal, we can only inspect a finite number of digits of a number. In order to check if a real number x equals to 0.00..., we would have to decide equality for an infinite number of digits. Assuming that we would have such an algorithm, we could in particular decide the halting problem.

When assuming lem, the unit interval is as "compact as it could be" (HoTT Book, 2013:p. 478). Hence, we can provide an inhabitant of the following type:

$$\mathsf{EQ0} :\equiv \prod_{x:\mathbb{R}} (x < 0) + (x > 0) + \mathsf{Id}(x, 0)$$

With this lemma, the above bisection method can be formalized in HoTT and an inhabitant of IV can be presented. Since the inhabitant of $\mathsf{EQ0}$ contains lem, the proof of IV will also contain lem. Since the term lem cannot be eliminated, we will not be able to produce a canonical form of the proof of IV, which means that we cannot extract a method that produces a $x : \mathbb{R}$ such that $f(x) = 0$. So while we cannot in general give a root of f, we can still claim that it is *true* that there is a root (note that approximate versions of the algorithm will of course work in HoTT, but we wanted to prove the existence of an exact solution).

We conclude that when we assume the law of excluded middle, we can establish the *truth* of a proposition, but will not have *computational content* in the proof. This is how it should be and everything we can ask for. Crucially, the assumption of LEM only destroys the constructive character of proofs that contain lem, whereas the other proofs remain unaffected. Hence, it is possible to discern between constructive proofs and classical proofs inside the same theory.

6.2 A Unified Theory of Meaning

Now that we have seen how we can employ the law of excluded middle in intuitionistic type theory, we want to explore the repercussions of such a procedure for our understanding of mathematics. We will try to give a theory of meaning for mathematical expressions that takes serious the insistence on giving constructive proofs while allowing for the application of the law of excluded middle. We will need to define what we consider the reference of both constructive and non-constructive proofs and lay down what we consider the meaning of the logical constants.

Martin-Löf's "syntactico-semantical" approach to meaning theory nicely characterizes the special status of canonical forms. However, a "contentual language" is an oxymoron from a classical point of view: A language is comprised of linguistic expressions and might only refer to some content outside of the language. This "outside" is what intuitionists discard as metaphysical speech, but we need some sort of determinate mathematical reality to make coherent the postulation of the LEM. Hence, we will try to keep form and meaning apart and understand type theory as an uninterpreted formal language (which is of course very apt to capture mathematics). We will still incorporate the intuitionistic account to the meaning of the logical constants and attributing a special status to canonical forms in our approach.

Note that the opposite of the syntactico-semantical approach does not consist in reducing the meaning of type theory to some set-theoretic model, such as Kripke frames as a semantics for the BHK meaning explanations. Rather, we only want to maintain that mathematical objects exist independently of the theory of types (we completely ignore the problems that result from stating the existence of abstract objects in this work). We can only hope that the types and terms we postulate in our type theory do refer to the mathematical objects under investigation.

We have to do the balancing act of combining traditional theories of reference with an inferentialistic view on the meaning of the logical constants and the special status of canonical forms. The following exposition should not be seen as a definite account, but rather as an outline of which questions need to be answered to provide a semantics for liberal constructivism.

Propositions and truth

The first decision that we need to make is how we define "proposition" — do we adopt the intuitionistic account of identifying a proposition with the set of its (canonical) proofs, or do we consider propositions to be simply truth-apt sentences? Since the LEM only makes sense for the latter understanding, we will adopt the traditional understanding of propositions as truth-apt sentences. We consider a proposition to be reflected with multiple types in the system, namely all those types that intuitively capture the proposition. For example, "$P(x)$ holds for some x" is represented both with $\Sigma_{x:A}P(x)$ and $\|\Sigma_{x:A}P(x)\|$. We can then establish the following truth-condition for a proposition:

> A proposition P is true iff a corresponding type A is inhabited, $a : A$.

Since a is not required to be canonical, we may employ the law of excluded middle in establishing the truth of P. The constructively valid propositions are those whose proofs can be reduced to canonical forms:

> A proposition P is constructively valid iff a corresponding type A is inhabited, $a : A$, and $\mathtt{a} \Rightarrow \mathtt{v}$ for some canonical form $\mathtt{v}$.

A proposition is constructively valid if a type corresponding to that proposition can be proved, and that proof evinces computational content. With this distinction between validity and constructive validity, we have at hand a very simple theory of reference. Going back to Frege and Carnap, the reference of a proposition is its truth-value. Hence, the reference of a proposition is true if an associated typed can be shown to be inhabited. For example, the intermediate value theorem has been proved above by means of the LEM, and we can establish the following reference pattern:

$$\mathsf{IV} \hookrightarrow \mathit{true}$$

In our theory of reference, we want to single out the canonical forms since they give direct access to the objects under investigation, as we have sketched in Section 4.4. We will incorporate Martin-Löf's meaning explanations by considering canonical forms as references of non-canonical terms and types. The reference of a canonical form in turn lies outside of the formal theory. For example, the reference of $\mathtt{add(s(s(0))s(s(0)))}$ is its canonical form

$\mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{s}(0))))$. We then maintain that this is still only a linguistic expression that has a reference lying outside of the language, namely in our intuitive conception of the natural number 4. This yields the following reference pattern.

$$\mathsf{add}(\mathsf{s}(\mathsf{s}(0))\mathsf{s}(\mathsf{s}(0))) \hookrightarrow \mathsf{s}(\mathsf{s}(\mathsf{s}(\mathsf{s}(0)))) \hookrightarrow 4$$

The natural number symbol $\mathtt{N}$ has an intended interpretation, but we leave it just as that: A symbol that refers to the structure that we have in mind when talking about the natural numbers. Types of HoTT are hence interpreted via the theory of expressions as follows:

$$\mathbb{N} \hookrightarrow \mathtt{N} \hookrightarrow \textit{Natural numbers}$$

So far for a theory of meaning for liberal constructivism. Since we asserted a determinate mathematical universe and interpreted the expressions with respect to that universe, it is coherent to use the LEM. Our theory takes serious the intuitionistic insistence on the special status of canonical forms, since we consider the reference of non-canonical expressions to be canonical forms. Canonical forms in turn have a reference in the mathematical universe and thereby give a direct connection to the mathematical objects under investigation.

The meaning of the logical constants

We not only have to specify the meaning of propositions and mathematical objects, but also of the constituent parts of mathematical sentences, namely the logical connectives. Hellman (1989) has argued that the logical symbols in intuitionistic logic differ in meaning from the ones used in classical logic, in particular, he argues that the existential quantifier $\exists$ and Σ are distinct in meaning. If we want to give a unified theory of meaning, we have to commit to one meaning of the logical constants.

Since the operation of propositional truncation introduced in Section 2.6 allows for removing the computational content of a type, it seems expedient to only retain the intuitionistic logical symbols and introduce classical quantifiers not as primitive symbols, but as abbreviations for truncated types. For example, the existential quantifier can be defined as follows:

$$\exists x.P(x) :\equiv \| \sum_{x:A} P(x) \|$$

The above type will be a mere proposition and will not contain any constructive content. Similarly, the classical conception of a disjunction can be recovered by:

$$P \vee Q :\equiv \|P + Q\|$$

Since $P \vee Q$ is propositionally equal to either $\mathbf{1}$ or $\mathbf{0}$, we cannot retain any information of which disjunct holds from a proof of $P \vee Q$.

If we adopt this convention, the demarcation line between classically and constructively valid proofs is not drawn by using different logical constants, but solely by the structure of the proofs. If we only have one inhabitant of a type, the proposition represented by the type is true.

This account seems to suggest that we have to subscribe to the inferentialistic understanding of the logical constants. The meaning of the symbols is understood in terms of how they are used, i.e., by the inference rules, and not by the conditions under which they are true. It needs to be investigated whether this leads to tensions with the referential theory stated above: The logical constants are defined in terms of constructability, whereas the meaning of the propositions is specified in terms of truth-conditions with respect to a mathematical universe — this discrepancy needs to be resolved. We want our semantic theory to be compositional in the sense that the meaning of the logical constants determines the meaning of more complex statements, if this can be maintained with the above expositions is unclear. Maybe, we need to delineate between referring and non-referring expressions in HoTT: Mathematical expressions refer to mathematical objects, whereas the meaning of the logical symbols lies inside the formal system. As an upshot, this would make apparent the "contentlessness" of logic. Additionally, we could employ proof-theoretic concepts such as harmony to justify the validity of the rules of type theory.

6.3 Drawbacks

The above exposition raises multiple concerns. In this section we want to present some breaking points of the theory, which are partially inherited from the incongruities of Martin-Löf's meaning explanations.

Propositional equality as equality of reference

In Section 2.5 we have sketched how propositional equality can be thought of as equality of reference, since from a structuralist point of view, isomorphic objects can be identified and propositional equality identifies precisely all structurally equal types. Complementary, we have characterized judgmental equality as equality of sense or presentation. This distinction has already been made by Per Martin-Löf and gives a fruitful way to think about both equalities.

If we adequately define a type $\mathbb{M}$ as containing an initial element and a successor element, we can establish that $\mathsf{Id}_{\mathcal{U}}(\mathbb{N}, \mathbb{M})$ is inhabited, i.e., that the type $\mathbb{M}$ is identical to the natural numbers type $\mathbb{N}$. It is then natural for a structuralist to claim that the reference of $\mathbb{M}$ is the natural numbers as well:

$$\mathbb{M} \hookrightarrow \textit{Natural numbers}$$

However, if $\mathbb{M}$ is not defined to be judgmentally equal to $\mathbb{N}$, we cannot establish the above denotation by means of translating $\mathbb{M}$ in the theory of expressions and evaluating it to canonical form — $\mathbb{M}$ is only propositionally equal to $\mathbb{N}$, but not judgmentally. This seems to suggest that we should take into account propositional equality when investigating if two expressions are coreferential.

This might also solve some other inherent problem in Martin-Löf's meaning explanations, which we have investigated in depth in Chapter 5: In intensional type theory, not every inhabitant of the identity type can be reduced to a canonical form. For example, if we have $p : \mathsf{Id}_A(a, b)$ different from both refl_a and refl_b, then we cannot assign a reference to p by evaluating it to normal form. However, p might be *propositionally* equal to refl_a witnessed by an inhabitant of $\mathsf{Id}_{\mathsf{Id}_A(a,b)}(p, \mathsf{refl}_a)$. In this case, we might consider p to be coreferential with refl_a. This might be unfolded in an adaption of Martin-Löf's meaning explanations that does justify the characteristics of identity in intensional type theory.

Another appeal of extending the canonical form relation to propositional equality could be the homotopy canonicity theorem, which we have introduced in Section 4.3.2: For simple types such as the natural numbers, we can prove that terms containing the univalence axiom are propositionally equal to a canonical numeral. This allows for assigning a reference to a term that is without reference according to Martin-Löf's meaning explanations. Further investigation in the extent and relevance of homotopy canonicity seems auspicious.

We conclude that a semantic theory for HoTT needs to take into account propositional equality. An appropriate incorporation might also solve the issue of non-canonical identity proofs.

Valid inference rules

If we postulate the existence of a term such as lem, we break the criterion for constructive validity since we introduce a term in the system that cannot be reduced to a canonical form. We argued informally that it is coherent to postulate the existence of lem — but are there other axioms that can or should be postulated? How can we delineate which terms are sensible and which are not?

Separating mathematics and logic

One latent problem that we had to deal with throughout the thesis is that logic and mathematics are conflated in the intuitionistic tradition. We have treated the set of natural numbers in the same way as a logical proposition, and a single number as a proof. Conversely, we have regarded proofs as first-class mathematical objects and have interpreted propositions as types which are richer in structure than truth-apt sentences.

This unification of proofs and objects is what makes HoTT so expressive and allows for an elegant formalization of homotopy theory, but is also philosophically challenging. In our semantic theory, we have distinguished between logical and mathematical expressions, which seems expedient since according to the orthodox view, logic is supposed to be without content, whereas mathematical expressions refer to some abstract mathematical entities. Identity is commonly regarded to be logical[1], but we were not able to give an account of identity in intensional type theory that makes it apparent that it is a logical constant. We will only be able to maintain that the logical part of HoTT does not refer if we can make sense of identity without the homotopy interpretation. Otherwise, identity is hardly without content since it is interpreted spatially.

Is it possible to uphold a systematic distinction between logic and mathematics? Is such a distinction even expedient? If the answer to these questions is no, we might have to revise our understanding of the relation between mathematics and logic.

6.4 Conclusions

We have seen in Section 6.1 that HoTT gives a neat way to distinguish between reasoning that yields constructive content and reasoning that only establishes the truth of a statement. The conviction that constructive reasoning is preferable to, but also compatible with classical reasoning is held by most mathematicians, and we hope to have presented a scaffolding for a semantic theory that does justice to this approach in Section 6.2. We have argued that the logical constants should be understood constructively in HoTT, but propositions can be given truth-conditions with respect to a determinate mathematical universe outside of the formal system. The special status of canonical forms stemming from Martin-Löf's meaning explanations have been incorporated in our semantical theory. There is still more work do be done in incorporating propositional equality. In particular, the distinctively structuralist aspect of HoTT with univalence has not been appreciated adequately in the semantic theory.

We hope that this gives a convincing argument that intuitionistic type theory is not committed to a meaning theory in the spirit of Prawitz and Dummett, even though more work is necessary to see if the verificationistic origins of intuitionistic type theory can be combined nonchalantly with a realist view on mathematics. It has to be investigated if some advantages of the intuitionistic account to epistemology can be saved and incorporated into the philosophy of mathematics of liberal constructivism.

In this thesis, we have not appreciated the subtleties of constructive mathematics. There are other principles that constructivists consider problematic apart from LEM, such as the limited principle of omniscience or the axiom of choice (HoTT Book (2013):Sect. 3.8 gives an overview of the status of AOC in HoTT, as some version of it does hold in HoTT). Often, multiple

1 Which is, of course, not without controversy, see most prominently Quine (1986).

constructive accounts of a concept compete. For example, there are several formulations of continuity that do satisfy the intermediate value theorem. Investigating the relations of these different conceptions seems vital to give a unified foundations for mathematics.

If these worries can be solved, we are able to put forward a philosophy of mathematics in support of liberal constructivism, which respects the special status of constructive proofs without surrendering the classically valid part of mathematics. Sundholm portrays the situation of mathematics in the beginning of last century as follows:

> [...] the foundation of mathematics were faced with the two horns of a dilemma: either we retain classical logic in a mathematical object-language, but give up hope for meaning explanations, or we insist on retaining a contentual language with meaning explanations, but have to jettison classical logic.
>
> — Sundholm (2012):p. xviii–xix

HoTT retains the computational content of constructive proofs while still allowing for the use of classical logic where necessary. The challenge is now to incorporate this into a coherent philosophy of mathematics. If this succeeds, the juxtaposition of constructive and classical reasoning can be resolved at last.

Index

Bibliography

Angere, Staffan (2017). “Identity and intensionality in Univalent Foundations and philosophy”. In: *Synthese*, pp. 1–41.

Atten, Mark van (2017). “The Development of Intuitionistic Logic”. In: *The Stanford Encyclopedia of Philosophy*. Winter 2017. Accessed July 29, 2019. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/entries/intuitionistic-logic-development/.

Awodey, Steve (1996). “Structure in Mathematics and Logic: A Categorical Perspective”. In: *Philosophia Mathematica* 4(3), pp. 209–237.

Awodey, Steve (2014). “Structuralism, Invariance, and Univalence”. In: *Philosophia Mathematica* 22(1), pp. 1–11.

Awodey, Steve (2018). “Univalence as a principle of logic”. In: *Indagationes Mathematicae* 29(6), pp. 1497–1510.

Awodey, Steve and Andrej Bauer (2004). “Propositions as [types]”. In: *Journal of logic and computation* 14(4), pp. 447–471.

Awodey, Steve and Michael Warren (2009). “Homotopy theoretic models of identity types”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 146(1), pp. 45–55.

Beeson, Michael J. (1980). *Foundations of Constructive Mathematics*. Springer.

Benacerraf, Paul (1965). “What numbers could not be”. In: *The Philosophical Review* 74(1), pp. 47–73.

Bentzen, Bruno (forthcoming). “On different ways of being equal”. In: *Synthese*.

Berger, Ulrich and Helmut Schwichtenberg (1995). “The greatest common divisor: a case study for program extraction from classical proofs”. In: *International Workshop on Types for Proofs and Programs*. Springer, pp. 36–46.

Billinge, Helen (2003). “Did Bishop have a philosophy of mathematics?” In: *Philosophia Mathematica* 11(2), pp. 176–194.

Bishop, Errett (1967). *Foundations of constructive analysis*. Reprinted in Bishop (2012). McGraw-Hill.

Bishop, Errett (1970). “Mathematics as a Numerical Language”. In: *Intuitionism and Proof Theory: Proceedings of the Summer Conference at Buffalo N.Y. 1968*. Vol. 60. Studies in Logic and the Foundations of Mathematics. Elsevier, pp. 53–71.

Bishop, Errett (1973). “Schizophrenia in contemporary mathematics”. In: *Errett Bishop: Reflections on Him and His Research*. Vol. 39. Conte. American Mathematical Society, pp. 1–32.
Bishop, Errett (2012). *Foundations of constructive analysis*. Ishi Press.
Brunerie, Guillaume (2018). “The James construction and $\pi_4(\mathbb{S}^3)$ in homotopy type theory”. In: *Journal of Automated Reasoning, Special Issue on Homotopy Type Theory and Univalent Foundations*, pp. 1–30.
Church, Alonzo (1940). “A formulation of the simple theory of types”. In: *The journal of symbolic logic* 5(2), pp. 56–68.
Coquand, Thierry and Gérard Huet (1986). “The calculus of constructions”. PhD thesis. Institut national de recherche en informatique et en automatique.
Curry, Haskell (1934). “Functionality in Combinatory Logic”. In: *Proceedings of the National Academy of Sciences* 20(11), pp. 584–590.
Davies, Edward Brian (2003). “Quantum mechanics does not require the continuity of space”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 34(2), pp. 319–328.
Davies, Edward Brian (2005). “A defence of mathematical pluralism”. In: *Philosophia Mathematica* 13(3), pp. 252–276.
Dummett, Michael (1991). *The logical basis of metaphysics*. Harvard University Press.
Dybjer, Peter (2012). “Program testing and the meaning explanations of intuitionistic type theory”. In: *Epistemology versus Ontology*. Springer, pp. 215–241.
Escardó, Martín Hötzel (2018). “A self-contained, brief and complete formulation of Voevodsky’s Univalence Axiom”. In: *arXiv:1803.02294*. Accessed 29 July 2019. URL: https://arxiv.org/abs/1803.02294.
Feferman, Solomon (1996). “Kreisel’s ’Unwinding Program’”. In: *Kreiseliana. About and Around Georg Kreisel*. A K Peters, pp. 247–273.
Fine, Kit (2014). “Truth-Maker Semantics for Intuitionistic Logic”. In: *Journal of Philosophical Logic* 43(2-3), pp. 549–577.
Frege, Gottlob (1892). “Über Sinn und Bedeutung”. In: *Zeitschrift für Philosophie und philosophische Kritik* 100(1), pp. 25–50.
Gentzen, Gerhard (1935). “Untersuchungen über das logische Schließen. I”. In: *Mathematische Zeitschrift* 39(1), pp. 176–210.
Gödel, Kurt (1958). “Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes”. In: *Dialectica* 12(3-4), pp. 280–287.
Goodman, Nicolas D (1983). “Reflections on Bishop’s philosophy of mathematics”. In: *The Mathematical Intelligencer* 5(3), pp. 61–68.
Granström, Johan Georg (2008). “Reference and Computation in Intuitionistic Type Theory”. PhD thesis. Uppsala University.
Granström, Johan Georg (2011). “Reference and Computation”. In: *Treatise on Intuitionistic Type Theory*. Springer, pp. 77–106.
Hedberg, Michael (1998). “A coherence theorem for Martin-Löf’s type theory”. In: *Journal of Functional Programming* 8(4), pp. 413–436.

Hellman, Geoffrey (1989). "Never Say "Never"!: On the Communication Problem Between Intuitionism and Classicism". In: *Philosophical Topics* 17(2), pp. 47–67.

Hellman, Geoffrey (1993a). "Constructive mathematics and quantum mechanics: unbounded operators and the spectral theorem". In: *Journal of Philosophical Logic* 22(3), pp. 221–248.

Hellman, Geoffrey (1993b). "Gleason's theorem is not constructively provable". In: *Journal of Philosophical Logic* 22(2), pp. 193–203.

Heyting, Arend (1931). "Die Intuitionistische Grundlegung der Mathematik". In: *Erkenntnis* 2(1), pp. 106–115.

Heyting, Arend (1934). *Mathematische Grundlagenforschung, Intuitionismus, Beweistheorie*. Springer.

Heyting, Arend (1956). *Intuitionism: an introduction*. North-Holland.

Hofmann, Martin and Thomas Streicher (1998). "The groupoid interpretation of type theory". In: *Twenty-five years of constructive type theory (Venice, 1995)*. Vol. 36. Oxford Logic Guides. Oxford University Press, pp. 83–111.

HoTT Book (2013). "Homotopy type theory". In: *Univalent Foundations Program, Institute for Advanced Study at Princeton*. Version: first-edition-1104-gd2c8502. URL: https://homotopytypetheory.org/book/.

Howard, William Alvin (1980). "The formulae-as-types notion of construction". In: *To HB Curry: essays on combinatory logic, lambda calculus and formalism* 44, pp. 479–490.

Klev, Ansten (2017). "The justification of identity elimination in Martin-Löf's type theory". In: *Topoi*, pp. 1–14.

Kohlenbach, Ulrich and Paulo Oliva (2003). "Proof mining: a systematic way of analyzing proofs in mathematics". In: *Mathematical logic and algebra* 242, pp. 147–175.

Kreisel, Georg (1951). "On the interpretation of non-finitist proofs—Part I". In: *The Journal of Symbolic Logic* 16(4), pp. 241–267.

Kreisel, Georg (1952). "On the interpretation of non-finitist proofs–Part II". In: *The Journal of Symbolic Logic* 17(1), pp. 43–58.

Kreisel, Georg (1966). "Foundations of Intuitionistic Logic". In: *Logic, Methodology and Philosophy of Science*. Vol. 44. Studies in Logic and the Foundations of Mathematics. Elsevier, pp. 198–210.

Ladyman, James and Stuart Presnell (2014). *A Primer on Homotopy Type Theory Part 1: The Formal Type Theory*. Accessed July 29, 2019. URL: http://philsci-archive.pitt.edu/11157/.

Ladyman, James and Stuart Presnell (2015). "Identity in Homotopy Type Theory, Part I: The Justification of Path Induction". In: *Philosophia Mathematica* 23(3), pp. 386–406.

Ladyman, James and Stuart Presnell (2016a). "Does Homotopy Type Theory provide a foundation for mathematics?" In: *The British Journal for the Philosophy of Science* 69(2), pp. 377–420.

Ladyman, James and Stuart Presnell (2016b). "Universes and Univalence in Homotopy Type Theory". In: *Review of Symbolic Logic*, pp. 1–30.

Ladyman, James and Stuart Presnell (2017). “Identity in Homotopy Type Theory: Part II, The Conceptual and Philosophical Status of Identity in HoTT”. In: *Philosophia Mathematica* 25(2), pp. 210–245.
Linnebo, Øystein and Richard Pettigrew (2011). “Category Theory as an Autonomous Foundation”. In: *Philosophia Mathematica* 19(3), pp. 227–254.
Martin-Löf, Per (1975). “An intuitionistic theory of types: Predicative part”. In: *Studies in Logic and the Foundations of Mathematics*. Vol. 80. Elsevier, pp. 73–118.
Martin-Löf, Per (1982). “Constructive Mathematics and Computer Programming”. In: *Studies in Logic and the Foundations of Mathematics* 104, pp. 153–175.
Martin-Löf, Per (1984). *Intuitionistic type theory*. Bibliopolis.
Martin-Löf, Per (1987). “Truth of a Proposition, Evidence of a Judgement, Validity of a Proof”. In: *Synthese* 73(3), pp. 407–420.
Martin-Löf, Per (1996). “On the Meanings of the Logical Constants and the Justifications of the Logical Laws”. In: *Nordic Journal of Philosophical Logic* 1(1), pp. 11–60.
Martin-Löf, Per (1998). “Truth and knowability: On the principles C and K of Michael Dummett”. In: *Truth in mathematics*, pp. 105–114.
McLarty, Colin (1988). “Defining sets as sets of points of spaces”. In: *Journal of Philosophical Logic* 17(1), pp. 75–90.
McLarty, Colin (1993). “Numbers Can Be Just What They Have To”. In: *Noûs* 27(4), pp. 487–498.
Peregrin, Jaroslav (2018). “Intensionality in Mathematics”. In: *Truth, Existence and Explanation*. Springer, pp. 57–70.
Petrakis, Iosif (2018). *Bishop’s constructivism in foundations andpractice of mathematics*. Accessed: July 29, 2019. URL: http://www.math.lmu.de/~petrakis/FMV.pdf.
Pfenning, Frank (2001). “Intensionality, Extensionality, and Proof Irrelevance in Modal Type Theory”. In: *Proceedings of the 16th Annual IEEE Symposium on Logic in Computer Science*. IEEE Computer Society, pp. 221–230.
Prawitz, Dag (1973). “Towards a foundation of a general proof theory”. In: *Studies in Logic and the Foundations of Mathematics*. Vol. 74. Elsevier, pp. 225–250.
Prawitz, Dag (1974). “On the Idea of a General Proof Theory”. In: *Synthese* 27(1/2), pp. 63–77.
Prawitz, Dag (1980). “Intuitionistic logic: a philosophical challenge”. In: *Logic and philosophy/Logique et philosophie*. Springer, pp. 1–10.
Prawitz, Dag (2006). “Meaning Approached via Proofs”. In: *Synthese* 148(3), pp. 507–524.
Quine, Willard Van Orman (1986). *Philosophy of logic*. Harvard University Press.
Schroeder-Heister, Peter (2006). “Validity concepts in proof-theoretic semantics”. In: *Synthese* 148(3), pp. 525–571.

Schroeder-Heister, Peter (2018). "Proof-Theoretic Semantics". In: *The Stanford Encyclopedia of Philosophy*. Spring 2018. Accessed July 29, 2019. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/spr2018/entries/proof-theoretic-semantics/.

Scott, Dana (1970). "Constructive Validity". In: *Symposium on Automatic Demonstration*. Springer, pp. 237–275.

Sommaruga, Giovanni (2000). *History and Philosophy of Constructive Type Theory*. Springer.

Streicher, Thomas (2015). "How Intensional Is Homotopy Type Theory?" In: *Extended Abstracts Fall 2013*. Springer, pp. 105–110.

Sundholm, Göran (1983). "Constructions, proofs and the meaning of logical constants". In: *Journal of Philosophical Logic* 12(2), pp. 151–172.

Sundholm, Göran (2012). "On the philosophical work of Per Martin-Löf". In: *Epistemology versus Ontology. Essays on the Philosophy and Foundations of Mathematics in Honour of Per Martin-Löf*, pp. xvii–xxiv.

Tait, William Walker (1983). "Against intuitionism: constructive mathematics is part of classical mathematics". In: *Journal of Philosophical Logic* 12(2), pp. 173–195.

Tait, William Walker (1994). "The Law of Excluded Middle and the Axiom of Choice". In: *Mathematics and Mind*. Oxford University Press, pp. 45–70.

Troelstra, Anne Sjerp and Dirk van Dalen (1988). "Constructivism in mathematics". In: *Studies in Logic and the Foundations of Mathematics* 121.

Tsementzis, Dimitrios (2016). "Univalence, Foundations and Philosophy". PhD thesis. Princeton University.

Tsementzis, Dimitrios (2017a). "Univalent foundations as structuralist foundations". In: *Synthese* 194(9), pp. 3583–3617.

Tsementzis, Dimitrios (2017b). "What is a Higher-Level Set?" In: *Philosophia Mathematica* 26(1), pp. 59–83.

Tsementzis, Dimitrios (2019). "A Meaning Explanation for HoTT". In: *Synthese*, pp. 1–30.

Tsementzis, Dimitrios and Hans Halvorson (2018). "Foundations and Philosophy". In: *Philosophers' Imprint* 18(10), pp. 1–15.

Voevodsky, Vladimir (2006). *A very short note on the homotopy λ-calculus*. Accessed July 29, 2019. URL: https://www.math.ias.edu/~vladimir/Site3/Univalent_Foundations_files/Hlambda_short_current.pdf.

Walsh, Patrick (2017). "Categorical harmony and path induction". In: *The Review of Symbolic Logic* 10(2), pp. 301–321.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit eigenständig und ohne fremde Hilfe angefertigt habe. Textpassagen, die wörtlich oder dem Sinn nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ort, Datum

Unterschrift