

Models of univalence in cubical sets

An introduction to the proof of univalence and a review of recent literature

Supervisor:

Prof. Dr. Dominique Devriese
Vrije Universiteit Brussel, Department
of Computer Science, Software
Languages Lab

Co-supervisor:

Prof. Dr. Wouter Castryck
Katholieke Universiteit Leuven,
Department of Mathematics, Algebra
Section

Mentor: Andreas Nuyts
Katholieke Universiteit Leuven,
Department of Computer Science,
IMEC-DistriNet

Reader: Prof. Dr. Frank Piessens
Katholieke Universiteit Leuven,
Department of Computer Science,
IMEC-DistriNet

Willem VANHULLE

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Mathematics

Academic year 2018-2019

Preface

I chose this topic for my master thesis because I wanted to know if the implementation of the axiom of univalence in cubical sets can make the life of a mathematician easier. And whether it could make writing, reading and exchanging proofs in obscure but still interesting mathematical domains more efficient.

This text was written with the help of many people. First of all, I would like to thank everyone who read drafts of this text thoroughly and gave very useful feedback: Andreas, Dominique, Frank and Wouter. Because I do not have a clue of the quality of my writing and these friendly people have a lot of experience, their reading of my drafts and feedback helped me a lot. I also had a lot of curious friends such as Jonathan, Michael, Dieter, Ben, Zhiquan. They listened to me when I tried to explain to them what I was reading or writing. Sometimes they asked questions which helped me even more.

Summary

The univalence axiom states that equivalent spaces can be identified and has many interesting consequences such as the ability to do mathematics up to isomorphism as claimed by [VAC⁺13]. If an axiom is consistent and simple enough, its addition to the theory is a compromise mathematicians are willing to make in exchange for beautiful results that depend on it. The consistency of the univalence axiom relative to ZFC was proven by giving a model in simplicial sets [KL12]. This was a first step in the recognition of the univalence axiom among mathematicians. However, this model in simplicial sets made use of the axiom of choice which made it unconstructive and unsuitable for a constructive implementation, see section 2.3.

Cubical type theory is an intuitionistic type theory that can model the univalence axiom, not as an axiom, but as a valid and provable theorem. This type theory will be explained in detail in chapter 3. Its proof in cubical theory has only recently been completed in [CCHM16]. Meanwhile, several updated proofs have been constructed in [SHC⁺18, MV18] which will be discussed in section 3.4. By the constructiveness of cubical type theory, a proof gives an explicit map between given equalities. The map can be used to rewrite proofs of theorems in homotopy theory in the more basic constructive language of cubical sets. Understanding constructive proofs of univalence may also help to study or understand applications or consequences of the univalence axiom which will be surveyed (including one new application) in section 3.5. Throughout this text, as much recent literature will be covered and in section 3.5.4 and section 3.6.3 possible future work is discussed.

Glossary

$\Gamma \vdash \dots, \Delta \vdash \dots$ Contexts

$\mathbb{G}\text{blue}$ Font for types

$A = B$ Equality of types

$X \equiv Y$ Equality by definition

$A \simeq B$ Equivalence of types

\mathcal{U} Universe of types

Set Category of sets

\mathcal{C}, \mathcal{D} Categories

\mathcal{F}, \mathcal{G} Functors

(i/r) A substitution of i by r

Notes

<input type="checkbox"/> https://arxiv.org/pdf/1803.02294.pdf	14
<input type="checkbox"/> Which one should I use	34
<input type="checkbox"/> Work out.	37
<input type="checkbox"/> Is the context right?	40
<input type="checkbox"/> I don't get example on Huber2016, p. 101	42
<input type="checkbox"/> add information of implementations of higher inductive types	53
<input type="checkbox"/> What about the relational semantics, non-fibrant types?	58
<input type="checkbox"/> Which parts of [Shi18] are interesting? This remark was also made by other learners of homotopy type theory such as [Boo18]	59

Contents

1	Introduction	1
1.1	Foundations of mathematics	1
1.1.1	Constructivism	1
1.1.2	Type theory	2
1.2	Intuitionistic type theory	2
1.2.1	Judgments and contexts	2
1.2.2	Informal type theory	4
1.2.3	Natural deduction	4
1.2.4	Propositions as types	5
1.2.5	Universes	7
1.2.6	Dependent types	7
1.2.7	Equality	8
1.2.8	The intensional identity type	10
1.3	The univalence axiom	13
1.3.1	Levels above sets	15
2	Interpretations and models	17
2.1	Groupoid interpretation	17
2.2	Homotopy type theory	18
2.3	Simplicial model	18
2.4	Categorical semantics	20
2.5	Presheaves	22
3	Cubical type theory	29
3.1	The face lattice	29
3.2	Manipulating contexts	32
3.3	Adding operations	35
3.3.1	The composition operation	36
3.3.2	The path type	37
3.3.3	The filling operation	40
3.3.4	The <code>glue</code> type	41
3.4	Proof of univalence	44
3.4.1	Statement of theorem	44
3.4.2	History of the proof	44
3.4.3	Contractibility of equivalence singletons	44
3.4.4	Conclusion of the proof	47
3.4.5	Univalence with topoi	48

3.5	Applications of the proof of univalence	48
3.5.1	Isomorphism invariant algebra	48
3.5.2	Generic datatypes	52
3.5.3	Formalizing algebraic topology	52
3.5.4	Future applications	54
3.6	Alternative cubical models of univalence	55
3.6.1	Nominal models	56
3.6.2	Cartesian computational type theory	56
3.6.3	Open challenges	58
	Index	60

Chapter 1

Introduction

1.1 Foundations of mathematics

Around the beginning of the twentieth century, mathematicians were working on foundations of mathematics and formalized the building blocks of mathematics: sets, propositions and proofs. Important examples of developments are the first inductive definition of the natural numbers which appeared in [Pea79], the concept of types invented around the time of [Rus03] and intuitionistic logic [Hey30]. Because these developments, especially intuitionism and type theory are important for the rest of this text and quite different from mathematics, this chapter will devote some time to explaining both.

1.1.1 Constructivism

The foundational work of the end 19th century by mathematicians such as Cantor's sets and ordinals was very abstract and vague, even to other mathematicians. Constructivism or intuitionism is the name for the philosophical view of that time that questions weird abstract foundational work and was made popular by [Bro05]. According to [Bro05], mathematics should be the result of mental human activity rather than objective discovery. This philosophy was formalized into a formal system based on intuitionism, called intuitionistic logic [Hey30].

An example of the problems that intuitionistic logic deals with, is giving proofs with computational content in the form of explicit constructions. It is for example not enough in intuitionistic logic to say that something exists or does not exist and this will not be accepted as a fact for an intuitionistic logician. These undesirable facts are instead stated as axioms such as the principle of excluded middle or the axiom of choice. One example of a theorem that traditionally is given using the principle of excluded middle, is mentioned in most introduction to constructive mathematics and also in the introduction of [Pal14]:

Theorem 1.1.1. *There are irrational numbers a and b such that a^b is a rational number.*

Throughout the course of the last century, using induction, types, intuitionistic logic and other developments in the foundations of mathematics and the theory of computation, intuitionistic type theory [Ml75] was established. Recently, the field of type theory became very popular among mathematicians with the arrival of more computing power. For example, many proof assistants are based on (intuitionistic) type theory:

Coq, NuPRL, Agda. These are programming languages that can be used to write down proofs.

On top of expressions and definitions, a user of a proof assistant can formulate complex theorems and proofs. These formalized proofs can be used to investigate proofs of real-life theorems from mathematical domains such as algebra or topology that are hard to write down by hand such as the Odd Order Theorem:

Theorem 1.1.2. *If G is a finite group and there exists an $n \geq 0$ such that $|G| = 2n + 1$, then G is solvable.*

This theorem was conjectured in 1911 and proven much later in [FT63]. It was one of the first proofs in group theory that were hundreds of pages long. The proof was formalized in the proof assistant Coq. The implementation of the formal proof ended up being at least (or only) 5 times as long [GAA⁺13], section 6. The implementation of such a large proof required the implementation of new features and re-usable libraries in the proof assistant Coq. Subsequently, new developments should be easier. Proponents claim that proof assistants will become more widely used in the short term.

1.1.2 Type theory

Type theory, in general, is different from set theory. Set theory has two main layers: sets and propositions about elements of sets. In type theory, the analogues of these two layers are called *types* and *terms* of types. The main difference between set theory and type theory is that terms are always accompanied by their type. Types may be empty, but when they have a term, they are called *inhabited*. Terms do not exist on their own. Without a type annotation, terms do not have any meaning, while in set theory it is perfectly possible to speak about a mathematical object on its own. For example, in type theory there is the type of groups denoted by `Group` and a group G would be denoted by the notation $G : \text{Group}$. Although this difference may seem very superficial, it implies that set membership is not a logical proposition any more.

Designers of formal systems and proof assistants based on type theory are confronted with questions about the properties of their systems. For example, certain expressions written in the system may or may not normalize or other expressions may be invalid in the system but not according to a mathematician using the system. These are the topics that this text will try to cover although some applications may be mentioned in section 3.5. An in-depth overview of the history of type theory can be found in [Coq13] or [Con11] and [Con15].

1.2 Intuitionistic type theory

Now that some useful applications of type theory have been given, this section will explain in more detail what intuitionistic type theory [Ml75] is about.

1.2.1 Judgments and contexts

A *judgment* is a statement in (intuitionistic) type theory about types and terms. Judgments can be philosophically seen as a constructive act of knowledge but formally, they come in different forms:

- A (starting with upper case) is just the statement that A is a type.
- $A = B$ means that A and B are equal types, can also be seen as an identity type, see def. 1.2.5.
- $a : A$ (starting with lower case) means that a is a term of type A , also called the *membership judgment*.

Judgements are always accompanied by transformation rules that describe how several judgements can be combined into one new judgement. These rules, called *typing rules*, can be seen as rules of computation. This is an important difference with set theory because in set theory the only constructions available are not much more than sets, tuples and propositions. But intuitionistic type theory is rooted in intuitionism and constructivism (see section 1.1.1) which requires a constructive approach to computation.

Definition 1.2.1. *The following types of typing rules are required for every type:*

- *The introduction of a new type and how the type is referenced once it is introduced is denoted with a formation rule. These rules are important for the readability of the judgements.*
- *How terms of the type being introduced are constructed is denoted by a typing rule. For example, there are two ways to introduce a term of the type of natural numbers: the zero element or its successor. These rules for introducing terms are called constructors.*
- *To compute with terms and types, we also need rules that describe how we can apply terms and types to other terms and types. These rules are called eliminators.*
- *We now gave rules that allow us to build complex expressions but to compute with these expressions of types and terms we need computation rules. These determine how an eliminator can act on a constructor.*

This kind of structural distinction between rules according to their constructive effect on the formal system is not really present in set theory. In set theory there is not much more than set comprehension. However, certain set-based theories such as category theory also have a more structural rule-based approach.

To make the bookkeeping of judgments and the above rules easier, *contexts* are used. Contexts contain the judgments that are valid in the type theory according to the above rules but are for some reason deemed not important enough by the type-theoretical proof writer to mention explicitly. Their usage corresponds to the usage of scopes in programming languages. Traditionally a context is denoted by Γ or Δ and stands for a list $a_1 : A_1, \dots, a_n : A_n$ of judgments $a_i : A_i$ where a_i is a term of the type A_i . The empty context is denoted by $()$. In later chapters, contexts do not always take this form (see def. 3.2.2). Judgments can be added to or taken from contexts:

- Judgments $a : A$ may depend on the judgments that are present in a context which is denoted by the expression $\Gamma \vdash a : A$. In some formal type theories, all judgments are required to have a context. When a judgment does not depend on any other judgment, this is denoted by $() \vdash a : A$ or in other words, $a : A$ depends on the empty context.

- A context Γ may be extended with any judgments $x : T$ if $\Gamma \vdash x : T$ is a valid expression in a process called *context extension*. Contexts may also be shortened by removing judgments and moving them towards the right. This is what happens for example in the introduction rule for the product type stating that if $\Gamma, x : A \vdash b : B$ is a valid judgment, then there is a term of the product type $\Gamma \vdash \lambda(x : A).b : \prod_{x:A} B$.

1.2.2 Informal type theory

The book [VAC⁺13] introduced a difference between two kinds of type theory: informal and formal. Informal type theory was introduced to make it easier for people accustomed to classical set-based mathematics to understand type theory. The usage of contexts in informal type theory is implicit and natural language is alternated with new judgements. This practice is clearly illustrated in the the first chapters of [VAC⁺13].

In *formal type theory*, for example as explained in [VAC⁺13], A.2 or in implementations in proofs assistants, context dependency is explicit because natural language is considered ambiguous. This was also one of the reasons type theory was introduced in the first place: to eliminate mathematical paradoxes formulated in natural language. There are two very famous proof assistants that have a different approach to dealing with contexts. On the one side there is Coq [ADH⁺19] which treats contexts as special objects that can be manipulated using typing rules, see def. 1.2.1. On the other hand there is Agda [ACD⁺19] which treats context dependency as a functional dependency which makes Agda less of a proof assistant than Coq.

However, in all proof assistants there are ways to leave parts of the context implicit and force the implementation to infer the missing parts. This concept brings back some of the benefits of informal type theory and is called *implicit arguments* but may return the problems of ambiguity as in informal type theory.

1.2.3 Natural deduction

Throughout this text, formal type theory will be studied as much as possible. Because context dependency is explicit in formal type theory, it is possible to write very rigorous and explicit sequences of derivations in proofs. These derivations apply rules for the interaction of terms, types and contexts, also called *derivations*, see def. 1.2.1. In this text and many other texts, there is a special notation for derivations that is inspired by natural deduction. The method of *natural deduction* was originally a formal approach to logic that helped to state the rules of the proving game very succinctly. It was a reaction to the less informative Hilbert-style proofs and was defined for the first time in [Gen35]. Proofs in the style of natural deduction visually resemble how proofs are constructed in the head of a mathematician, see fig. 1.1.

Understanding this notation is important for understanding the literature on the subject. Most texts on type theory introduce the typing rules at the beginning with the help of natural deduction. The notation allows for a very succinct description and introduction. For example, the main reference text on cubical type theory introduces type theory on one page in fig. 1 of [CCHM16].

$$\begin{array}{c}
\frac{[A] \quad [\neg A]}{\perp} \Rightarrow E \\
\frac{\perp}{\neg\neg A} \Rightarrow I \\
\frac{\neg\neg A}{A \Rightarrow \neg\neg A} \Rightarrow I
\end{array}$$

Figure 1.1: An example of the proof of the proposition $A \Rightarrow \neg\neg A$ formalized with natural deduction. In each step a new rule is applied and the resulting transformed judgement is denoted below a horizontal line.

1.2.4 Propositions as types

Functions, which are terms of function types, are the most fundamental types and building blocks of type theory.

Example 1.2.2. *The typing rule for elimination or computation of function types, see def. 1.2.1 is as follows:*

$$\frac{\Gamma \vdash f : A \rightarrow B \quad \Gamma \vdash a : A}{\Gamma \vdash b : B}$$

This typing rule does not make use of set-theoretic concepts such as relations but just reflects how functions are used or computed with.

Function types were historically the first types to have an interpretation in intuitionistic logic [Hey30] which was discovered in [Cur34]. More precisely, implications in intuitionistic logic correspond to terms of function types in type theory. This means that a constructive proof of the proposition $A \Rightarrow B$ corresponds to giving a function in type theory which is a term f of the function type $A \rightarrow B$, written with a membership judgment as $f : A \rightarrow B$.

Types that can be constructed with the function type are called *simple types* but other types also have parallels with logic:

Example 1.2.3. *The sum type of two types A and B , denoted by $A + B$, corresponds to the “or” operation, formally denoted by the symbol \vee from intuitionistic logic. In proofs that make use of the logical proposition $A \vee B$, there are two branches: one branch that proves the case were we assume A and the other branch proves the case were B is assumed. The terms of a sum $A + B$ are also constructed in two ways: either such a term is left $a : A + B$ for $a : A$ or right $b : A + B$ for $b : B$. A function $A + B \rightarrow C$ for C some type is defined by case analysis, stating a result for values of both injections.*

This interpretation can be generalized somewhat to interpreting constructive propositions as types and their proofs to terms of types. But it remains just an approximate interpretation between two completely foundational theories, not an isomorphism or bijection. This interpretation can however be helpful to formalize constructive mathematics and was applied in the first proof assistant Automath [BG70] (see fig. 1.2). Multiple people worked (independently) on making this interpretation precise and it is often named after (some of) these people. Besides being called an interpretation, it is also called the propositions-as-types relation or *Curry-Howard correspondence*. There are several long introductions available to the history of this correspondence such as [Bib19].



Figure 1.2: Nicolaas Govert de Bruijn (1918-2012) in Oberwolfach was the creator of one of the first proof assistants. Less-known, he started to treat proofs in full mathematical logic as objects with computational content. Picture taken from [Jac17].

$$\begin{aligned}
 \llbracket P \Rightarrow Q \rrbracket &\equiv \llbracket P \rrbracket \rightarrow \llbracket Q \rrbracket \\
 \llbracket P \wedge Q \rrbracket &\equiv \llbracket P \rrbracket \times \llbracket Q \rrbracket \\
 \llbracket P \vee Q \rrbracket &\equiv \llbracket P \rrbracket + \llbracket Q \rrbracket \\
 \llbracket \forall x : A. P \rrbracket &\equiv \prod_{x:A} \llbracket P(x) \rrbracket \\
 \llbracket \exists x : A. P \rrbracket &\equiv \sum_{x:A} \llbracket P(x) \rrbracket \\
 \llbracket \perp \rrbracket &\equiv 0 \text{ (empty type)} \\
 \llbracket \top \rrbracket &\equiv 1 \text{ (singleton type)}
 \end{aligned}$$

Figure 1.3: The Curry-Howard correspondence can be illustrated by defining a map Ψ that takes an intuitionistic proposition P and extracts its evidence $\llbracket P \rrbracket$ as a type. Ψ can be defined for the full first-order fragment of constructive mathematics, based on [Alt17], p.3.

Consequences of the correspondence

Certain versions of type theory do have good computational properties such as being strongly normalizing: computations terminate. The main requirement for being strongly normalizing is not having axioms such as excluded middle and other versions of the axiom of choice. In other words, the type theory can only contain type theoretic interpretations of concepts of constructive mathematics, see section 1.1.1. The Curry-Howard correspondence then implies for such intuitionistic type theories that terms double as constructive and computable proofs. Which means that these terms can be converted into programs, a process called *program extraction* and implemented in many proof assistants. Examples of such implementations are:

- The proof assistant Agda is a functional programming language based on intuitionistic type theory def. 1.2.9 in which programs double as mathematical proofs [ACD⁺19]. Program extraction is possible for well-behaved parts of the theory but it is not the main purpose.
- Idris is a programming language related to Agda and compiles to imperative languages, see [BCdMH18].
- The proof assistant Coq [ADH⁺19] has program extraction as one of its main applications, see for example [PCZF⁺18]. This particular extraction mechanism does have some limitations, for example in the context of the Theorem of Algebra [CFL05].

1.2.5 Universes

You may have asked yourself the question already how a type is defined or how type theory deals with the usual self-referencing paradoxes of set theory. Actually, it is done by making the assumption that there exist a upside-down tower of types and each type somewhere belongs in that tower. The levels of the tower are *indexed* by ordinals or natural numbers and it is assumed that every type belongs to some level of the tower. Assuming that the indexing set is the natural numbers, the levels are denoted by \mathcal{U}_i for $i \in \mathbb{N}$. The levels are called *universes* and are cumulative, $T : \mathcal{U}_i : \mathcal{U}_{i+1} \Rightarrow T : \mathcal{U}_{i+1}$.

In informal type theory, this tower is implicitly used and it is assumed that every type T belongs to some universe, denoted by \mathcal{U} or *Type*. The universes in the tower can be considered as types, with types as terms. An important consequence of this is that every type is also a term of some universe.

In formal type theory implementations such as Agda, the tower of universe levels can and has to be manipulated explicitly. This allows for formalizations of complex foundational mathematics. This task is also necessary because the assignment of universe levels is a non-trivial task that cannot be automated without introducing inconsistencies, see fig. 1.4.

1.2.6 Dependent types

In traditional logic, propositions state properties about elements of sets. Such a property could for example state: given a group, there is a group element g with infinite order. One could say that this property depends on the presence of an element g , in a sense it is

$$\begin{aligned} & _ \simeq _ : \forall (A : \text{Set } \ell) (B : \text{Set } \ell') \rightarrow \text{Set } (\ell\text{-max } \ell \ \ell') \\ & A \simeq B = \Sigma [f \in (A \rightarrow B)] (\text{isEquiv } f) \end{aligned}$$

Figure 1.4: Equivalences will be introduced in def. 1.3.1. The definition of equivalences in [MV18] is defined by explicitly referencing the universe levels of the types A and B , respectively denoted by ℓ and ℓ' . The universe level of the type of equivalences is the sum $\ell\text{-max } \ell \ \ell'$ to prevent paradoxes arising from using $A \simeq B$ in the same context as other expressions using A or B .

a dependent property depending on a group element g . Properties can be translated to type theory through the Curry-Howard correspondence as types such that types, being interpreted as propositions, may depend on terms of other types. In this case, they are called *dependent types*.

Example 1.2.4. *Let A and B be types belonging to the hierarchy of universes in a consistent way, this means that their presence does not introduce any paradoxes. Formally speaking, it is enough to require that A and B belong to the same universe type: $A, B : \mathcal{U}$. Next to the simple types, the following types are basic building blocks for intuitionistic type theory:*

- *The type of finite products is $A \times B$ and has as terms the tuples of the form (a, b) where $a \in A$ and $b \in B$. When the second term b depends on the first term a , this is denoted as $b(a) : B(A)$. The type of all terms that have this dependency is called the dependent sum and denoted by $\sum_{a:A} B(a)$ or in implementations simply as $\sum A \ B$. The formal typing rules for the dependent sum as in def. 1.2.1 can be seen in [VAC⁺13], A.2.5.*
- *The dual type of the dependent sum is the dependent product and is denoted by $\prod_{a:A} B(a)$. The difference with the dependent sum is that the first component a of its term becomes an argument and only the dependent part is returned. This means that the dependent product can be viewed as a generalization of the function type and a dependent product $f \in \prod_{a:A} B(a)$ is written as a dependent function $f : (a : A) \rightarrow B(a)$. The typing rules (see def. 1.2.1) for the dependent product are formally stated in [VAC⁺13], A.2.4.*

When A only has two terms, both the dependent product and the dependent sum are equivalent with the finite product type $B \times B$. Using these two dependent types it is possible to build other dependent types. See fig. 1.5 for an example of a custom dependent types used in practice. When working in set-theoretic models, the definition of a dependent type can be stated very precise in terms of sets, see def. 2.4.8.

1.2.7 Equality

Until now we only used equality by definition which is denoted in this text by \equiv . Equality by definition can be seen as an equality up to substitution of definitions. The *intensional identity type*, also called judgmental or propositional equality, was only fairly recently added to type theory. Its definition, see def. 1.2.5, was given inductively, similar to the inductive definition of the natural numbers (see fig. 1.6).

```

notZero :  $\mathbb{N} \rightarrow \text{Set}$ 
notZero n =  $\Sigma \mathbb{N} (\lambda m \rightarrow (n \equiv (\text{suc } m)))$ 

 $\mathbb{N}_0$  :  $\text{Set } \_$ 
 $\mathbb{N}_0$  =  $\Sigma \mathbb{N} (\lambda n \rightarrow \text{notZero } n)$ 

```

Figure 1.5: The type of non-zero natural numbers \mathbb{N}_0 can be defined in Agda using a dependent sum type indexed by the natural numbers. Terms of the type \mathbb{N}_0 are tuples of a natural number and a proof that it is non-zero.



Figure 1.6: Per-Martin L  f, born in 1942, is a Swedish logician. He was the first to give a good definition of equality in type theory with his intensional identity type [Ml75]. Picture taken from [Eur13].

When the intensional identity type is added to the types of functions and dependent products, it gives a completely new theory, called type theory (see def. 1.2.9). This theory had good computational properties such as the *normalizability* of terms and types. This implies it is suitable for implementation of proof assistants and this contributed to its popularity. Type theory is also called MLTT and is very close to the implementation of proof assistant Agda [ACD⁺19] and variants of MLTT such as CIC, used in proof assistant Coq [ADH⁺19]. For now, for understanding the intensional identity type and the univalence axiom, the differences are however not that important. There are for example formalizations of the univalence axiom and applications of it in both assistants. For simplicity, I will pretend that MLTT, CIC and type theory with the intensional identity type are all the same and denote them by *type theory*. But before it is possible to speak of type theory, a description of the intensional identity type is necessary.

1.2.8 The intensional identity type

The intensional identity type is not defined as relation with tuples, opposed to the traditional equivalence relations of mathematics. The intensional identity type is a type with its own typing rules which will be stated explicitly in def. 1.2.5. The intensional identity type can be applied to terms of the same type. Because two terms that have an inhabited intensional identity type, satisfy all the same properties, the intensional identity type conveys the traditional meaning of judgmental equality $=$ between objects in mathematics. This property, actually saying that equals can be substituted, is also called *indiscernability of identicals*. As a result, we can treat the intensional identity type as the equality we know which is casually denoted by “=”:

Definition 1.2.5. *Given a type X and terms $a, b : X$, the intensional identity type between a and b is written as $a = b$, $a =_X b$ (or sometimes $\text{Id}_X(a, b)$). The intensional identity type satisfies prop. 1.2.6 and section 3.3.2.*

So what are the constructors and eliminators of the intensional identity type that teach us how its terms interact with their surroundings?

Property 1.2.6. *There is only one constructor called reflexivity or in short refl . Given a term $a : X$ it returns the term $\text{refl}(a)$ of the intensional identity type $a = a$.*

The definition of the eliminator for the intensional identity type is one of the most complex expressions in type theory because it looks so self-explanatory or simple but is not. This is because the proof of equality is now a term of the intensional identity type and there can be multiple terms (or equalities).

Property 1.2.7. *The eliminator for the intensional identity type is also called path induction or J . It states: given a dependent product type, also called a family*

$$C : \prod_{x, y : A} (x =_A y) \rightarrow \mathcal{U}$$

representing a predicate depending on terms (and proofs of identity), a function

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

that can be considered as the base step in an inductive construction, there is a function

$$f : \prod_{x,y:A} \prod_{p:x=y} C(x,y,p)$$

such that

$$f(x,x,\text{refl}_x) \equiv c(x).$$

It tells us that to prove a property for all terms x, y and equalities $p : x = y$ between them, it suffices to consider all the cases where x is definitionally equal to y and where the term of the intensional identity type under consideration is $\text{refl}_x : x = x$.

The intensional identity type is reflexive by construction. Using the definition of the eliminator, it is possible to prove the remaining properties of the usual identity [VAC⁺13], chapter 2.

Lemma 1.2.8. *For every type A and every $x, y, z : A$:*

- *The intensional identity type is symmetric. There is a function $(x = y) \rightarrow (y = x)$ denoted by $p \mapsto p^{-1}$, such that $\text{refl}_x \equiv \text{refl}_x$ for each $x : A$. The term p^{-1} is called the inverse of p .*
- *The intensional identity type is also transitive. This means that there is a function $(x = y) \rightarrow (y = z) \rightarrow (x = z)$ written $p \mapsto q \mapsto p \star q$ such that $\text{refl}_x \star \text{refl}_x \equiv \text{refl}_x$. The expression $p \star q$ is called the concatenation of p and q .*
- *The intensional identity type supports indiscernability of identicals. If P is a type family over A or in other words, there is a dependent function $P : A \rightarrow \mathcal{U}$ and suppose there is a term $p : x =_A y$, then there is a function, called the transport $p_\star : P(x) \rightarrow P(y)$. When the family P is important, p_\star is sometimes also denoted by $\text{transport}^P(p, -) : P(x) \rightarrow P(y)$.*

These properties are however not the defining properties of the intensional identity type. This is what makes the intensional identity type special compared to the traditional notion of equality.

As mentioned earlier, the introduction of the intensional identity type was an important step in the establishment of (Per-Martin L  f) type theory. Having defined the intensional identity type in def. 1.2.5, it is possible to summarize the previous sections in a definition. Because there are many definitions of type theory, it is also a good occasion to fix on a particular definition.

Definition 1.2.9. *A theory about types is called a (computational) type theory if it includes typing rules as in def. 1.2.1 for the following types:*

- *The empty type and the unit type.*
- *The inductive type \mathbb{N} of natural numbers.*
- *The function, dependent product, sum types and their derivatives introduced in section 1.2.6.*
- *The intensional identity type from def. 1.2.5.*

- *The universes from section 1.2.5.*

This definition is however very simple and is treated much more in detail in introductions to type theory or Martin-Löf type theory such as [Pal14] or [VAC⁺13]. In this text, only aspects of type theory relevant to (applications of) the proof the univalence theorem will be mentioned.

Implications of equality as a type

With equality having been defined and implemented, we can translate the definitions of mathematical objects that use equality relations between elements of sets to the language of type theory. An example is the type theoretic definition of a group. One could try to define a group as a dependent sum. Take G to be the base type, e the neutral element, i the inversion and m the multiplication map. A first try at a definition of a group would then be $\sum_{G,e,i} m : (G \times G \rightarrow G)$, but this definition is not yet complete. The maps i and m have to satisfy associativity and other properties. These properties have to be encoded as types or terms. Let's say we have a proof that these properties hold, called α , then we just extend the first definition with this proof. However, we might ask ourselves the following question: if two groups G and G' have different multiplication maps, they are definitely not equal. But what if they have different proofs of associativity, α and α' ? This may happen because the identity between terms of G may have more than one term. Are they still equal? They are not, at least according to the way we chose our base type G . By the definition of the intensional identity type on G it was possible to have different terms of the intensional identity type and accordingly different proofs of associativity α and α' .

Of course this is an undesirable consequence of the generality of the definition of the intensional identity type. In practice, we may not want to have more than one term or proof of equality between certain terms representing group elements. The difference in proofs is irrelevant to the structure and properties of the group. This motivates the following definitions:

Definition 1.2.10. • *An h-prop X is a type such that any two elements of it are equal and can be formally stated as*

$$\text{isProp}(X) \equiv \prod_{x,y:X} (x = y).$$

- *When the intensional identity type between any two terms of a type X is an h-prop, the type is called an h-set which is formally denoted by*

$$\text{isSet}(X) \equiv \prod_{x,y:A} \prod_{p,q:(x=y)} (p = q).$$

It can be proven that $\text{isSet}(X)$ is always a proposition.

The terminology of these concepts refers to the fact that they serve the same purpose as propositions and sets in intuitionistic logic. It can be proven that they have the same properties. In type theory the prefix is omitted, for example the type of all types that are h-props is denoted by Prop and the type of all types that are h-sets is denoted by Set , but for clarity, this text will always try to use the “h” prefix.



Figure 1.7: Vladimir Voevodsky (1966-2017), here seen in his office at Princeton, was an algebraist who found new ways to apply topology to algebraic geometry. When working on the simplicial model for a type theory that was inspired by algebraic topology, [Voe10] introduced the univalence axiom under the name of equivalence axiom. He was also one of the organizers of the Univalent Foundations project that resulted in a textbook for univalent type theory. The practical aspects of this project are discussed in [Bau13]. According to [Voe14], the term *univalence* comes from a Russian translation of [BV06] where the term “faithful functor” is translated as “univalent functor”. Vladimir Voevodsky said about his terminology himself: “Indeed these foundations seem to be faithful to the way in which I think about mathematical objects in my head.” Unfortunately, Vladimir Voevodsky suffered from alcohol addiction and died because of an aorta-aneurism [Gra17].

In the previous example, the type G can be chosen to be an h-set, that is $G : \mathbf{Set}$. Then the possibility of having multiple proofs of associativity is eliminated because the intensional identity type is always an h-prop. The definition of groups can be changed such that it only allows for G to be an h-set and we obtain an intuitive type-theoretical alternative to the traditional mathematical definition of groups. This means that all existing constructive proofs can be studied again in type theory. Although this may seem like a stupid idea to do everything again, it has proven to be useful in the context of big theorems in algebra that needed a lot of so-called proof-engineering [GAA⁺13].

1.3 The univalence axiom

The univalence axiom states that the type of equivalences between two types is equivalent with the intensional identity type between those types. Its addition to type theory gives a new theory, called *univalent type theory*, see def. 1.3.3, a new domain in mathematics that was considered very promising and still is. In this domain, concepts from topology, category theory and type theory are combined to study new ways of writing proofs and discover new results because of its topological interpretations, see section 2.2. The definition of the univalence axiom is based on the concept of equivalences. It is proven in [VAC⁺13], section 4.5, that there are at least three different but equivalent definitions of equivalence. The definition of equivalences used in this section and

in [PW14] corresponds with the definition `isContr` in [VAC⁺13]. In applications (see for examples section 3.5) it is often easier to use `biinv` which views equivalences as a kind of generalized homotopies.

Definition 1.3.1. *Let $X, Y : \mathcal{U}$ for some universe \mathcal{U} be types and $f : X \rightarrow Y$ an ordinary function.*

- *For each $y : Y$, the homotopy fiber of f over y is defined as*

$$f^{-1}(y) \equiv \sum_{x:X} (f(x) =_Y y).$$

- *A type X is contractible, if there is an $y : X$, called the center of contraction, such that the type $x = y$ is inhabited, there is a term contr_x , for all $x : X$. In other words $\text{isContr}(X) \equiv \sum_{y:X} \prod_{x:X} (x = y)$.*
- *The function f is called an equivalence between X and Y if there exists a term of the type*

$$\text{equiv}(f) \equiv \prod_{y:Y} \text{isContr}(f^{-1}(y)).$$

The type of all equivalences between the types X and Y is given by

$$X \simeq Y \equiv \sum_{f:X \rightarrow Y} \text{equiv}(f).$$

Given two types $X, Y : \mathcal{U}$ for some universe \mathcal{U} and a term $p : (X = Y)$ of the intensional identity type, it is quite straightforward to prove with the transport that there is a unique equivalence between the types $e : (X \simeq Y)$. This gives a function $\Phi_{X,Y}$ from the intensional identity type to the type of all equivalences.

Axiom 1.3.2 (Univalence axiom). *Given types $X, Y : \mathcal{U}$ for some universe \mathcal{U} , the map $\Phi_{X,Y} : (X = Y) \rightarrow (X \simeq Y)$ is an equivalence of types.*

/arxiv.org/pdf/1803.02244.pdf

It can be proven that all equivalences as defined in def. 1.3.1 are propositions. The univalence axiom, once stated and accepted, has to be a proposition because the property of being an equivalence is a proposition. The univalence axiom gives us ways to identify equivalent types. An algebraic topologist [Gra18b] said that it

... offers the hope that formalization and verification of today's mathematical knowledge may be achievable, relieving referees of articles of the tedious chore of checking the details of proofs for correctness, allowing them to focus on importance, originality, and clarity by exposition.

It was also mentioned in [Voe10] that the use of the univalence axiom in proof assistants such as Coq could be useful for the formalization of mathematics. This later resulted in an actual library [VAG⁺10] with code formalizing large portions of popular mathematics. Meanwhile, there are more introductions available to this approach to univalent type theory such as the course text [Esc19].

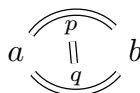


Figure 1.8: The type $\text{Id}_{\text{Id}_X(a,b)}(p, q)$ can be interpreted as a “second layer” of intensional identity types.

Definition 1.3.3. *A type theory (see def. 1.2.9) that includes the univalence axiom, defined in ax. 1.3.2, is called univalent or a univalent type theory. Because of its homotopic interpretation, see section 2.2, it is also called a homotopy type theory.*

Assume we have proven that h-sets have the same properties as traditional sets. Take two h-sets, or even two definitionally equal h-sets. In general, there are multiple bijections possible between these h-sets. Bijections between h-sets are the equivalences between h-sets. Simplifying the finer aspects of equivalences, the univalence axiom roughly gives a bijection between those equivalences and terms of the intensional identity type between h-sets. However, in case there are multiple bijections, we obtain multiple equalities and this is unsatisfactory. An h-set should only be equal to itself in one way. Furthermore, the universe in which those h-sets are contained cannot be an h-set. This demonstrates two facts that hold in univalent type theory, the type theory with the univalence axiom: not all types are h-sets and we may have multiple proofs of identity, the *uniqueness of identity proofs principle* (UIP) does not hold. So univalent type theory has some slightly counter-intuitive consequences for the classical mathematician that is used to working with (h-)sets and never really bothers too much with alternative interpretations of equality. The chapters on algebraic topology and category theory in [VAC⁺13] show what some other consequences are of accepting the univalence axiom for classical mathematical domains and rejecting the UIP principle.

1.3.1 Levels above sets

Evidence against and consequences of the UIP principle can be illustrated with the following example:

Example 1.3.4. *If $a, b : X$, there might be two terms $p, q : \text{Id}_X(a, b)$. Because the intensional identity type can be formed or introduced between any two terms of any type, there is also an intensional identity type $\text{Id}_{\text{Id}_X(a,b)}(p, q)$, see fig. 1.8.*

By the Curry-Howard correspondence (see section 1.2.4), this second layer contains the proofs of equality between proofs of equality. It is possible to continue this construction inductively and obtain a third layer and so on. One could ask the question whether this tower of equalities ever stops?

Previously, we have seen that the acceptance of the univalence axiom forces us to reject the UIP principle. For h-sets the tower contains at most two levels. But what about other types, types that are not h-sets?

There is a type such that the type of its equalities is not a set. For example the type of objects in a category is not an h-set, see [VAC⁺13], lemma 9.1.8. In general, there

exist constructions of types that prove for any positive integer n that there is a tower of equalities which is non-trivial with depth n . This proves that not all types are h-sets and also that the UIP principle is invalid in general types.

Chapter 2

Interpretations and models

The presence of a non-trivial hierarchy of equalities in type theory was somewhat unexpected and gave rise to a new field in foundations of mathematics that also had many applications [VAC⁺13]. Mathematicians who are working in foundations can try to find models of a theory they are working with. Models with category theory prove the relative consistency of the system in relation with classical foundational frameworks such as set theory, intuitionistic logic and optionally the axiom of choice. This was also the case in (univalent) type theory which was first modelled with groupoids in [HS98]. This section will explore some of the models that exist for univalent type theory.

2.1 Groupoid interpretation

Before the univalence axiom was discovered, one of the first models of type theory was done in the language of groupoids [HS98]. A *groupoid* is a concept of category theory that is inspired by the behavior of groups. A groupoid is a category, consisting of a collection of invertable morphisms between objects. Let X be any type. The model in groupoids of type theory worked by interpreting every term of $a, b : X$ as an object and the terms of the intensional identity type $\text{Id}_X(a, b)$ as invertable morphisms (see fig. 1.8 where the equalities are replaced by morphisms). The tower of equalities needed to be interpreted as well. This was done with a higher-dimensional groupoid, called an ∞ -groupoid or ω -groupoid, see the thesis [Lum10] for an extensive overview of the model in groupoids. An ∞ -groupoid contains morphisms between morphisms and so on. However, for the model to be a correct model, it also needs to satisfy all the properties that are satisfied by the tower of equalities in type theory. These properties of equalities are translated as *coherence laws* of the ∞ -groupoid and stated for example that the composition of two morphisms between morphisms of the ∞ -groupoid should be associative. But these coherence laws became very complex for higher levels of morphisms in the ∞ -groupoid and there was not a clear simple way to state them generically.

The tower of equalities inspired [Voe16] to define a hierarchy of levels of types above \mathbf{h} -sets. This hierarchy, that can be proven to be cumulative, is defined by induction on the integers, starting from -2 .

Definition 2.1.1. *A type X has \mathbf{h} -level -2 if it is contractible. It has \mathbf{h} -level $n + 1$ if for any term $a, b : X$, the intensional identity type $\text{Id}_X(a, b)$ has \mathbf{h} -level n .*

For example, h-props are by their definition on level -1 and h-sets are on level 0 . Now the more complicated types living “above” sets on higher h-levels can be referenced as well. Why are these more complicated? Take the example of the type *Group* and two groups, terms $G, H : \text{Group}$. In the groupoid model, the terms of intensional identity type $\text{Id}_{\text{Group}}(G, H)$ between two terms are modeled by the isomorphisms between the groups G and H . But the isomorphisms between two groups, in the category of groups, form a set. Assume the groupoid model correctly describes type theory, this implies that the type of groups lives above the level of sets and has h-level 1 . Something similar can be done for all types coming from categories, not only groups, by only considering the invertible morphisms, the core of the category. This can be interpreted as h-level 1 types corresponding to categories.

According to [Voe16] the h-level hierarchy can help to distinguish different levels of doing mathematics. On the bottom, there is element-level mathematics, the level corresponding to h-level -1 . Working with h-level -1 types corresponds to working with elements of sets in classical mathematics. There is also set-level mathematics, which is about working with h-level 0 types, the types that correspond to h-sets or sets in classical mathematics. On top of all of this, there is higher-level mathematics, the mathematics that is about type-theoretical analogues to categories, ∞ -groupoids or other higher-dimensional category theory.

2.2 Homotopy type theory

Besides the groupoid model from section 2.1, there are also other models. The model of type theory and more specifically univalent type theory, was motivated by the complexity of the ∞ -groupoid model and algebraic topology. In algebraic topology, the properties of topological, continuous spaces are studied with methods from group theory. In the homotopic interpretation of type theory, the type X is interpreted as a topological space and the intensional identity type $\text{Id}_X(a, b)$ is interpreted as all the continuous paths between points a, b in X considered as topological space. Because paths are reversible, the symmetry of the intensional identity type is consistent and the other properties of equality can be verified to be consistent as well. Let $p, q : \text{Id}_X(a, b)$ be two terms interpreted continuous paths (compare this with the construction in example 1.3.4). Terms $r, s : \text{Id}_{\text{Id}_X(a, b)}(p, q)$ are interpreted as two-dimensional paths, see fig. 2.1. This explains the naming *homotopy type theory* to univalent type theory, when interpreted using algebraic topology.

The definition of h-levels can be done in this interpretation as well. Here, a type is on h-level n if its analogue in the homotopy theoretic model, a topological space, has vanishing homotopy groups above order n . This topological model can be extended to all concepts in univalent type theory by giving interpretations to each concept in the language of homotopy theory, see table 2.1. So the study of types in univalent type theory corresponds to the study of higher order homotopy groups.

2.3 Simplicial model

The topological spaces used in the homotopic interpretation of univalent type theory are actually not really topological spaces but a kind of simplicial sets, Kan complexes

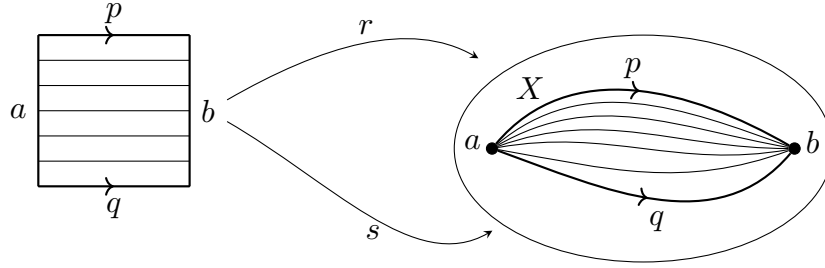


Figure 2.1: In the context of topological spaces and algebraic topology, a path between two curves p, q starting and ending at the same base points a, b is called a *homotopy of paths*. In this picture there are homotopies s and r . Formally, a homotopy of paths is in general a continuous function $F : [0, 1] \rightarrow X$ that satisfies $F(0, u) = a, F(1, u) = b, \forall u \in [0, 1]$ and $F(v, 1) = p(v), F(v, 0) = q(v), \forall v \in [0, 1]$. The study of homotopic paths leads to the study of fundamental or homotopy groups, see section 3.5.3.

Logic	Types	Homotopy
proposition	A	topological space
proof	$a : A$	point in space
$A \Rightarrow B$	$A \rightarrow B$	continuous functions
predicate	$B : A \rightarrow \mathcal{U}$, denoted $B(x)$	covering
conditional proof	$b(x) : B(x)$	section of a covering
\perp, \top	$0, 1$	\emptyset, \star
$A \vee B$	$A + B$	co-product
$A \wedge B$	$A \times B$	product space
$\exists_{x:A} B(x)$	$\sum_{(x:A)} B(x)$	total space of a covering
$\forall_{x:A} B(x)$	$\prod_{(x:A)} B(x)$	space of sections
equality	Id_A	path space $A^{[0,1]}$

Table 2.1: The homotopic interpretation of type theory as written in the introduction of [VAC⁺13] gives a comparison of the concepts of univalent type theory together with their interpretations in the topological model and intuitionistic logic, see section 1.2.4.



Figure 2.2: Martin Hofmann (1965-2018) was the founder of the groupoid interpretation [HS98] and popularized the usage of categorical semantics in type theory [Hof97]. In January 2018, he died in a snow storm on the mountain Nikko Shirane [Rod18].

with fibrations. See [Str06] and [Voe09] for how universes are modeled in simplicial sets. The simplicial set model was completed in [KL12] and proved to be very useful for later models of univalent type theory such as the one in [CCHM16]. The model in simplicial sets does however have the same homotopic properties, the homotopy of spaces stays the same whether we view them as topological spaces or Kan complexes, so it is a simplification that can be made. This is also called the *homotopy hypothesis* [SCS⁺18]. Because the model in simplicial sets was proven to be consistent, it also proved that univalent theory as a mathematical theory was as consistent as Zermelo-Fr ankel set theory with the axiom of choice. The usage of the axiom of choice made it non-constructive however, meaning that the translation of proofs in univalent type theory into the language of simplicial sets also became non-constructive. It was as well cumbersome to implement in computers, because simplicial sets and more specifically their building blocks, simplexes, do not have nice combinatorial properties. This lead to the development of cubical type theory, see chapter 3.

2.4 Categorical semantics

Both the groupoid and simplicial model turned out to have some problems. To properly deal with these problems, new constructive and easy to implement models have to be found. Models of type theory can be built with the framework of *categories with families* [Hof97], see fig. 2.2.

Categories with families can be seen as algebraic representations of type theory. The idea behind a category with families is to translate the types, terms and typing rules from type theory to a model in category theory.

Definition 2.4.1. *A category with families model is a category and a functor denoted by (Ctx, \mathcal{F}) that satisfies props. 2.4.2 and 2.4.4 to 2.4.7.*

Property 2.4.2. *The context category Ctx is a category whose objects represent contexts of the type theory. There is a terminal object $()$ in Ctx , called the empty context. This is an object such that for every other object $\Gamma \in Ctx$, there is a unique morphism $\Gamma \rightarrow ()$.*

The objects of this category are denoted with capital Greek symbols Δ, Γ and are simply called contexts, even though they only represent or model the contexts from type

theory. If the choice for Ctx is clear from the context, $\Gamma \vdash$ denotes that Γ is a context in the context category Ctx . The morphisms of the category Ctx are called *substitutions* or *context morphisms*.

The first thing we need is a way to capture the relation between terms and their types in set (and category) theory. This is done with a family of sets.

Definition 2.4.3. *The category of families of sets **Fam** has as objects pairs of the form (A, B) where A is a set and $B = (B_a \mid a \in A)$ is a A -indexed family of sets $B_a, a \in A$.*

A morphism between $(A, B) \rightarrow (A', B')$ is given by a pair (f, g) where $f : A \rightarrow A'$ is a function and g is an A -indexed family of functions such that $g_a : B_a \rightarrow B'_{fa}$.

Property 2.4.4. *Types and terms are modeled by a functor $\mathcal{F} = (\text{Ty}, \text{Tm})$ that goes from the opposite category Ctx^{op} to the category of families **Fam**.*

This functor takes a context Γ , returns a set denoted by $\text{Ty}(\Gamma)$, called the *types* in context Γ and a family of sets indexed by elements $A \in \text{Ty}(\Gamma)$, called the terms $\text{Tm}(\Gamma, A)$ of A . Because these category-theoretical types represent the real types, the dependency $A \in \text{Ty}(\Gamma)$ is often denoted by the judgement $\Gamma \vdash A$. Similarly, the elements $a \in \text{Tm}(\Gamma, A)$ are called *terms* of A within the context Γ , denoted by $\Gamma \vdash a : A$.

Take a substitution $\sigma : \Delta \rightarrow \Gamma$, then the morphism $\mathcal{F}(\sigma)$ is a morphism that acts on $\text{Ty}(\Gamma)$.

Property 2.4.5. *The action of a substitution on a type $A \in \text{Ty}(\Gamma)$ is written as $A\sigma$ and the action on terms a is written as $a\sigma$. Because \mathcal{F} is contravariant, $A\sigma \in \text{Ty}(\Delta)$ and $a\sigma \in \text{Tm}(\Delta, A\sigma)$. Let Γ' be yet another context. The substitutions on terms should satisfy the following composition laws: $(A\sigma)\tau = A(\sigma\tau)$ and $(a\sigma)\tau = a(\sigma\tau)$ for $\tau : \Gamma' \rightarrow \Delta$.*

Because contexts in type theory interact with other judgements, our category-theoretical model should be able to do this as well. For example, if Γ is a context and A is a type in Γ , there is a context $\Gamma.A$ called the *context extension*. Intuitively, this context has all the terms and types that are already in Γ and the type A , including the term $\Gamma \vdash q : A$ defined in prop. 2.4.6.

Property 2.4.6. *Given a context Γ and type $\Gamma \vdash A$, we postulate the existence of an object representing the context extension $\Gamma.A$, a morphism $p : \Gamma.A \rightarrow \Gamma$, and a term $q \in \text{Tm}(\Gamma.A, Ap)$, denoted by $\Gamma.A \vdash q : Ap$.*

Here the term q stands for the variable in A that was added the latest to $\Gamma.A$

Property 2.4.7. *The objects p and q should satisfy the universal extension property: for any context Δ, Γ , substitutions σ, τ and term a there is a substitution $(\sigma, a) : \Delta \rightarrow \Gamma.A$ such that*

- *The substitution σ is a partial inverse of p : $p(\sigma, a) = \sigma$.*
- *The substitution (σ, a) of the term q returns a : $q(\sigma, a) = a$.*
- *The substitution is invariant under composition of σ with other substitutions τ : $(\sigma, a)\tau = (\sigma\tau, a\tau)$.*

- The context morphism (p, q) is the identity on $\Gamma.A$.

There is an alternative presentation of the category with families model in [Ort19], section 3.1.

Given a category with families, we can formalize the informal description in section 1.2.6 with a very short definition:

Definition 2.4.8. *A type B is a dependent type depending on another type A within context $\Gamma.A$ if $\Gamma.A \vdash B$.*

If B is a dependent type, substitution of terms in B is done by lifting the morphism or substitution from the universal property of extensions where we take $\sigma = 1$ the identity substitution of contexts. We then obtain a morphism $(1, a) : \Gamma \rightarrow \Gamma.A$ that can be lifted with the Ty-functor. This lift is denoted by $[a] : \Gamma \rightarrow \Gamma.A$. The action of $[a]$ on B gives $B[a]$ which is by the contravariance of \mathcal{F} a type in the context Γ , so $\Gamma \vdash B[a]$.

2.5 Presheaves

In practice, a category with families is built with more fundamental concepts. A base category is chosen such that it possesses the properties that we desire from the contexts in the particular type theory we are modelling. With this base category, a category with families is built. For example, if we want the contexts of the type theory to be indexable by dimension variables, we have to choose a base category that has objects which represent dimensions.

Examples of presheaves

Definition 2.5.1. *Let \mathcal{C} be any category. A presheaf on \mathcal{C} is a contravariant functor from \mathcal{C} into the category of sets, **Set**. This is a mapping of objects and morphisms of categories that reserves arrow.*

- The category of all presheaves on base category \mathcal{C} has as objects the presheaves and as morphisms the natural transformations between presheaves. It is denoted by $Psh(\mathcal{C})$.
- Let I, J be two base objects the base category and $\Gamma \in Psh(\mathcal{C})$ a presheaf. The mapping of a morphism $f \in Hom_{\mathcal{C}}(J, I)$ under Γ is formally denoted by $\Gamma(f)$ but often just f and written as a right action. It is a morphism $\Gamma(f) : \Gamma(I) \rightarrow \Gamma(J) : \rho \mapsto \rho f$ where we say that ρf is the restriction of ρ by f .

Example 2.5.2. *If we take as base category \mathcal{C} a category with only one object \star , then the presheaves in $Psh(\mathcal{C})$ are contravariant functors from \star to **Set** mapping the object to any set.*

Example 2.5.3. *Let $\mathcal{C} = \{0, 1\}$ with two morphisms $f, g : 0 \rightarrow 1$.*

$$0 \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} 1$$

A presheaf Γ on \mathcal{C} consists of two sets $\Gamma(0), \Gamma(1)$ together with two functions $\Gamma(f), \Gamma(g) : \Gamma(1) \rightarrow \Gamma(0)$.

$$\Gamma(1) \begin{array}{c} \xrightarrow{\Gamma(f)} \\ \xleftarrow{\Gamma(g)} \end{array} \Gamma(0)$$

If we interpret $\Gamma(0)$ as the set of vertices and $\Gamma(1)$ as the set of edges, we have a directed graph. The map $\Gamma(f)$ gives the starting vertex of an edge, $\Gamma(g)$ gives the ending vertex.

Example 2.5.4. Take $\mathcal{C} = \{0, 1\}$, two morphisms $f, g : 0 \rightarrow 1$ as before and now a third morphism $h : 1 \rightarrow 0$ such that $h \circ f = h \circ g = \text{id}$. Then we interpret $\Gamma(0)$ again as a set of vertices and $\Gamma(1)$ as a set of nodes.

$$\Gamma(1) \begin{array}{c} \xrightarrow{\Gamma(f)} \\ \xleftarrow{\Gamma(h)} \\ \xleftarrow{\Gamma(g)} \end{array} \Gamma(0)$$

In this interpretation, we have that $\Gamma(f) \circ \Gamma(h) = \text{id}_{\Gamma(0)}$ and $\Gamma(g) \circ \Gamma(h) = \text{id}_{\Gamma(0)}$. This means that for each $v \in \Gamma(0)$, $\Gamma(h)(v)$ serves as an edge with both starting and ending vertex v . This means that from every presheaf, we can derive a unique reflexive graph.

More examples of presheaves will arise later or can be found in [Hof14]. There does exist a concept of *sheaves*. Presheaves can be turned into a sheaf in a topological space by a sheaving functor. This implies presheaves are a generalization of sheaves contrary to the terminology. More specifically, sheaves are a *reflective subcategory* of presheaves, see [S⁺11]. Another related concept is the concept of a sieve, which arises in the context of cubical type theory [Coq15] and [CCHM16].

The presheaf model

But first, let's explain how we can make a concrete category with families using presheaves. This explanation is based on [Hof97], section 4.1, page 45, [Hub16a], section 1.2 and [Ort19], section 3.2.

Definition 2.5.5. Given a small category \mathcal{C} (a category in which the morphisms form sets), called the base category, a presheaf model on \mathcal{C} is defined as the category of families $(Psh(\mathcal{C}), (Ty, Tm))$, where the individual terms are defined in props. 2.5.6, 2.5.8 and 2.5.9 and have a specific mechanism for dealing with contexts defined in prop. 2.5.10.

The first thing needed for the presheaf model to satisfy the definition of a category with families (see def. 2.4.1) is a category of contexts.

Property 2.5.6. In a presheaf model over a base category \mathcal{C} , we take the category of contexts to be the presheaf category $Psh(\mathcal{C})$. The empty context is the functor 1 mapping all the objects in the base category to the same set.

The resulting category then satisfies prop. 2.4.2.

We also need to give a functor from the category of presheaves to families of sets that satisfies prop. 2.4.4 for types and terms in a category with families. To define this, we first need to introduce another category.

Definition 2.5.7. *Given a functor $A : \mathcal{C} \rightarrow \mathbf{Set}$, the category of elements is the category denoted by $\int_{\mathcal{C}} A$:*

- *The objects are pairs (c, a) where c is an object of \mathcal{C} and $a \in A(c)$.*
- *Take two objects (c, a) and (c', a') . Morphisms between these objects are denoted by $(\phi \mid a)$ for $\phi : c \rightarrow c'$ and $a' = A(\phi)(a)$.*

If the functor A is contravariant, we define

$$\int_{\mathcal{C}} A = \left(\int_{\mathcal{C}^{\text{op}}} A \right)^{\text{op}}.$$

We can now construct the functor \mathcal{F} on the category of contexts from prop. 2.4.4. Let Γ be a context in the presheaf model, this means that Γ is a contravariant functor from the base category to the category of sets.

Property 2.5.8. *The interpretation of the types over a given context is denoted by $\text{Ty}(\Gamma)$ and defined as $\text{Psh}\left(\int_{\mathcal{C}} \Gamma\right)$.*

These presheaves form a set because the base category is small, so the morphisms form sets. This means that the presheaves can be used as the indexing set of a family of sets and the definition of Ty in prop. 2.5.8 makes sense.

Given a presheaf model, there is familiar terminology available from type theory. When $A \in \text{Ty}(\Gamma)$, A is called a type over context Γ or even use the type theoretical notation $\Gamma \vdash A$. But A is not a set of on its own or does not contain any terms. To work with types, it is necessary to look at the values of A (as a functor) which are written as $A\rho$ given $\rho \in \Gamma(I)$.

Property 2.5.9. *The interpretation of the terms over some context Γ and type A in Γ is denoted by $\text{Tm}(\Gamma, A)$ and defined as all the elements*

$$a \in \prod_{I \in \mathcal{C}, \rho \in \Gamma(I)} A(I, \rho)$$

such that $\forall f \in \text{Hom}_{\mathcal{C}}(J, I), (a\rho)f = a(\rho f)$. In other words, to every base object I , set element $\rho \in \Gamma(I)$ a set element $a\rho$ is chosen in the set $A(I, \rho)$ defined by the given presheaf A .

The definitions in props. 2.5.8 and 2.5.9 make sure that (Ty, Tm) satisfies the requirements of prop. 2.4.4 and is an appropriate functor as required in def. 2.4.1.

Again, one often pretends that a term a in a pre-sheaf model is an object on its own by identifying this interpretation with the object in type theory it stands for. However, to analyse the interpreted object deeper, we need to observe which values $a\rho$ it takes “above” an element $\rho \in \Gamma(I)$ which is how we defined above in the presheaf model.

The definition of a category of families also requires us to define extension of contexts for our presheaf model.

Property 2.5.10. *Given a context Γ and type $\Gamma \vdash A$, we define the context extension in the presheaf model as a contra-variant functor. Given an object $I \in \mathcal{C}$, it takes values*

$$(\Gamma.A)(I) = \{(\rho, u) \mid \rho \in \Gamma(I), u \in A(I, \rho)\}.$$

This new functor is denoted by $\Gamma.A$. Given $\rho \in \Gamma(I), u \in A(I, \rho)$ and a morphism $f \in \text{Hom}_{\mathcal{C}}(J, I)$, $(\rho, u)f$ is defined as $(\rho f, uf)$.

The map p in the presheaf model is defined by

$$p : \Gamma.A \rightarrow \Gamma : (\rho, u) \mapsto \rho.$$

In other words, this is just the projection forgetting the second argument. Similarly, the term $q \in \text{Tm}(\Gamma.A, A\rho)$ is defined in the presheaf model as a mapping (polymorphic for an implicit object I)

$$q : (I \in \mathcal{C}) \rightarrow I \times (\Gamma.A)(I) \rightarrow A(I, \rho) : (\rho, u) \mapsto u$$

that forgets the first argument.

The above construction gives an instance of the definition of a category with families (see def. 2.4.1): a presheaf model. This model is made using only the objects and morphisms in the base category and its presheaves. However, the exact representation depends on the choice of map for Ty and Tm . In very simple cases of base categories, the choice is limited. For example, the previously mentioned examples of graphs and sets become categories with families. This means that categories with families are things that really exist. When the base category is more complicated, the contexts and types behave very differently but there is a characterization of types in terms of context extensions given in lemma 2.5.12.

Definition 2.5.11. *An equivalence of categories between categories \mathcal{C} and \mathcal{D} is a pair of functors $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \rightarrow \mathcal{C}$ that satisfies the following properties:*

- *The composition of functors $(\mathcal{F} \circ \mathcal{G}) : \mathcal{D} \rightarrow \mathcal{D}$ is isomorphic through a natural transformation with the identity functor $I_{\mathcal{D}}$.*
- *There is a natural equivalence of functors $(\mathcal{G} \circ \mathcal{F}) \cong I_{\mathcal{C}}$.*

Lemma 2.5.12. *Take an arbitrary base category \mathcal{C} . If $\Gamma \in \text{Psh}(\mathcal{C})$, then there is an equivalence between the category of types $\text{Ty}(\Gamma)$ and the category*

$$U \equiv \{(\Delta, \sigma) \mid \Delta \in \text{Psh}(\mathcal{C}), \sigma \in \text{Hom}_{\text{Ctx}}(\Delta, \Gamma)\}$$

with as morphisms $\text{Hom}_U((\Delta_1, \sigma_1), (\Delta_2, \sigma_2))$ the context morphisms $\phi : \Delta_1 \rightarrow \Delta_2$ such that the following diagram commutes:

$$\begin{array}{ccc} \Delta_1 & \xrightarrow{\phi} & \Delta_2 \\ & \searrow \sigma_1 & \swarrow \sigma_2 \\ & \Gamma & \end{array}$$

Proof. Let the first functor for the equivalence of categories be given by

$$\mathcal{F} : \text{Ty}(\Gamma) \rightarrow U : A \mapsto (\Gamma.A, p).$$

The second functor, the candidate for an inverse of \mathcal{G} , is given by

$$\mathcal{G} : U \rightarrow \text{Ty}(\Gamma) : (\Delta, \sigma) \mapsto ((I, \rho) \mapsto \{s \in \Delta(I) \mid \sigma(s) = \rho\}).$$

I will now prove that the composite functor $\mathcal{F} \circ \mathcal{G}$ is isomorphic to the identity functor on U through a natural transformation. Take an object $(\Delta, \sigma) \in U$, we have to prove that

$$(\mathcal{F} \circ \mathcal{G})(\Delta, \sigma) \cong (\Delta, \sigma).$$

But by writing out the definition of \mathcal{G} , we have that $(\mathcal{F} \circ \mathcal{G})(\Delta, \sigma) = \mathcal{F}((I, \rho) \mapsto \{s \in \Delta(I) \mid \sigma(s) = \rho\})$ which is in turn equal to

$$(\Gamma.((I, \rho) \mapsto \{s \in \Delta(I) \mid \sigma(s) = \rho\}), p).$$

This can be further simplified to

$$((I \mapsto \{(\rho, s) \mid \rho \in \Gamma(I), s \in \Delta(I), \sigma(s) = \rho\}), p).$$

We can call the result of the computation

$$(\mathcal{F} \circ \mathcal{G})(\Delta, \sigma) \equiv (\Gamma.A, p).$$

To prove $(\Gamma.A, p) \cong (\Delta, \sigma)$, we have to prove that there is an isomorphism of contexts $\phi : \Delta \rightarrow \Gamma.A$ such that the following commutative diagram commutes:

$$\begin{array}{ccc} \Delta & \xrightarrow{\phi} & \Gamma.A \\ & \searrow \sigma & \swarrow p \\ & \Gamma & \end{array}$$

Because ϕ should be a morphism of contexts and morphisms of contexts are natural transformations of functors, we have to define ϕ by a natural transformation. A natural transformation is given by morphisms $\alpha_I : \Delta(I) \rightarrow \Gamma.A(I)$ for all $I \in \mathcal{C}$ that satisfies commutativity properties. Set $\alpha_I : s \mapsto (\sigma(s), s)$ for every $I \in \mathcal{C}$. Each α_I is a bijection with inverse $\alpha_I^{-1} : (\rho, s) \mapsto s$. It only remains to prove the commutativity properties.

We also have to prove the reverse direction. Take a type $A \in \text{Ty}(\Gamma)$. We have to prove that $(\mathcal{G} \circ \mathcal{F})(A) \cong A$ in the category of types, being presheaves on the category of elements. So take an element $(I, \rho) \in \int_{\mathcal{C}} \Gamma$. We have to prove that

$$((\mathcal{G} \circ \mathcal{F})(A))(I, \rho) \cong A(I, \rho)$$

as sets. The left-hand side is by definition

$$\{s \in (\Gamma.A) \mid p(s) = \rho\} = \{(\rho', u) \mid \rho' \in \Gamma(I), u \in A(I, \rho), \rho' = \rho\}$$

and this is in bijection with $A(I, \rho)$.

We have now proven the equivalence. □

Example 2.5.13. *If we take as base category \mathcal{C} the category with two objects and three morphisms as in example 2.5.4, we obtain a category with families with as contexts directed reflexive graphs. Take a context in this category with families which is a graph Γ . We will study the properties of types and terms in this context. A type in the context Γ is according to prop. 2.5.8 a functor A from the category of elements to the category of sets:*

- *Every vertex in $\{(0, v) \mid v \in \Gamma(0)\}$ is mapped to a set $A(0, v)$. Every edge in $\{(1, e) \mid e \in \Gamma(1)\}$ is mapped to a set $A(1, e)$.*
- *The morphism $(f \mid e)$ in the category of elements is mapped to a set map $A(f) : A(1, e) \rightarrow A(0, \Gamma(f)(e))$. We can do the same for the morphism g . Let $\Gamma(f)(e) = v$. f can be interpreted as sending an edge $e' \in A(1, e)$ in a new graph to an endpoint, a vertex v' which lies in a set $A(0, x)$ above the endpoint v of e in original graph Γ .*
- *The morphism $(h \mid v)$ is mapped to a set map $A(h) : A(0, v) \rightarrow A(1, \Gamma(h)(v))$ which takes a vertex v' and maps it to an edge $A(h)(v') = e'$.*
- *The composition property $A(h \circ f) = A(f) \circ A(h)$ implies that*

$$id = A(f \mid \Gamma(h)(v)) \circ A(h \mid v)$$

which is a map

$$A(0, v) \rightarrow A(0, \Gamma(f) \circ \Gamma(h)(v)) : v' \mapsto v'.$$

The same can be done with g . This can be interpreted as before that $A(h \mid v)(v')$ is an edge e' with identical start and end point v' .

This means that a type A is a refined reflexive graph of Γ with more vertices and edges. The edges and vertices of Γ' lie above those of the original graph Γ , see ???. Between the original graph and the new graph Γ' there is a morphism that maps all the new vertices and edges on the original ones. The types in this context can also be more indirectly characterized with the previous lemma as tuples of a graph Δ and a graph morphism $f : \Delta \rightarrow \Gamma$.

A term is according to prop. 2.5.9 an element of $a\rho$ in $A(I, \rho)$ that depends on the choice of $\rho \in \Gamma(I)$ and satisfies an extra condition:

$$\forall J \in \mathcal{C}, f \in \text{Hom}_{\mathcal{C}}(J, I), (a(I, \rho))\Gamma(f) = a(I, \rho\Gamma(f)).$$

- *If $I = 0$, an element $\rho \in \Gamma(0)$ is an edge in Γ . We can check how the maps $f, g : 0 \rightarrow 1$ act on terms. Assume for example that f (actually $\Gamma(f)$) is the map giving the source of edges. We can do the same for g . The condition on terms tells us that $(a\rho)f \equiv (a\rho)A(f \mid e)$ has to be equal to $a(\rho\Gamma(f))$. In other words, the source of $a\rho$ should be a vertex $a(\rho\Gamma(f))$ in the graph Γ' of the vertex set that is above the source ρf (formally $\rho\Gamma(f)$) of ρ .*
- *If $I = 1$, an element $\nu \in \Gamma(1)$ is a vertex in Γ' and $a\nu$ is a vertex above ρ . Assume that $\rho f = \nu$. The only interesting map that acts on ν is the map h (or at least $\Gamma(h)$). νh gives the reflexive edge going from ν to ν . $a\nu$ also has a reflexive edge $(a\nu)h$. Now the condition says that the lift of the lower reflexive edge $a(\nu h)$ is the same one as the reflexive edge on the lifted vertex $(a\nu)h$.*

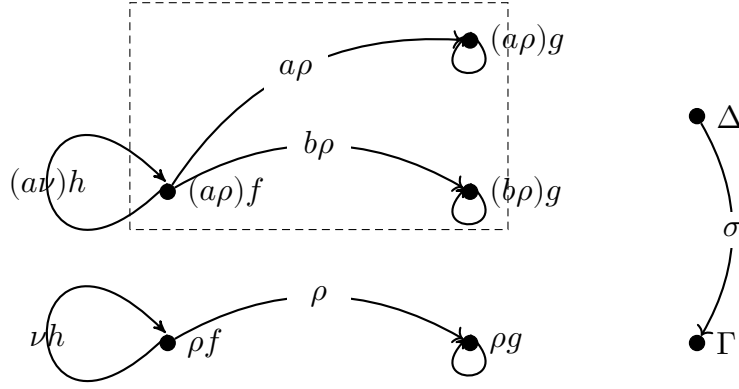


Figure 2.3: If b is another term, we can visualize the action of the morphisms f and g on the edges $a\rho$ and $b\rho$ described by the terms a and b . The term also acts on the node $\nu = \rho f$ and elevate it to higher nodes $a\nu$ and $b\nu$. This illustrates that this particular presheaf model as defined in def. 2.5.5 is a nested or dependent graph Γ' that lifts the original graph Γ . The rules that are required by props. 2.4.4 and 2.5.9 tell that the lifting of nodes and edges is compatible.

The presheaf model on a category with only two elements is in a sense the simplest example of a category with families there is. As soon more objects are added to the base category, the resulting category with families can be seen as a generalization of a reflexive directed graph.

A category with families supports a certain type if it is closed under the type rules of this type. It can be proven that every presheaf model supports all types from type theory, see [Hof97], [Nuy18], page 9 or [Hub16a], section 1.2. However, the traditional interpretation of the intensional identity type in the presheaf model does not satisfy the typing rules of the intensional identity type in (univalent) type theory, see [Hub16a], section 1.2.3. The traditional intensional identity type is a constrained version that satisfied an extra axiom, called *axiom K* stating:

Axiom 2.5.14 (Axiom K). *Let X be any type, if $\forall x : X$ and $p : (x = x)$, then $p = \text{refl}_x$.*

Because axiom K is related to the uniqueness of identity proofs principle (a relation that is covered in [VAC⁺13], section 7.2), it is in contradiction with the univalence axiom, see ax. 1.3.2. A presheaf model does not necessary support the intensional identity type from type theory, an identity type that supports multiple terms. In univalent type theory the identity type must have multiple terms and not satisfy ax. 2.5.14. This means that not any presheaf model is sufficient to model univalent type theory. To find a presheaf model for univalent type theory in category theory, we need to look further into variations on the presheaf model explained above.

Chapter 3

Cubical type theory

At this point we have discussed a general framework for making models of type theory in category theory, see section 2.4. However, remember from section 2.3 that we were actually looking for a very specific model. We would like a constructive model in category theory that allows us to prove the univalence axiom by generating terms of instances of it and using these terms in further computations. The model in cubical sets [BCH14] (and the slightly different one in [CCHM16]) was proven to be a model of univalent type theory, so the univalence axiom holds in it and it possible to generate terms of it. Actually, there are even more constructive models of the univalence axiom but in this text the focus will be on [CCHM16] and its recent extensions in [CHM18] and [MV18], see section 3.6.1 for other options.

The model with cubical sets can help us with computing with terms of the univalence axiom because it is easier to implement than other models and completely constructive. This means that all the evaluation the generated terms in this model does not get stuck and can always be simplified or computed into the primitives of the model. Proof of concept are the recent implementations of this model in the proof assistant Agda [MV18].

3.1 The face lattice

Cubical type theory is based on cubical sets which are built with cubes. The vertices of the cubes will describe the steps in applying a chain of (univalent) equalities as defined in def. 1.2.5. Computing with equalities in the cubical set model is done by computing with cubes, so we need to formulate precisely what a cube is. This will be done with an algebra of faces that was not yet present in [BCH14] but added in [CCHM16]. Here, faces are a generalization of the concept of vertices, edges and faces of 3-dimensional cubes to their higher-dimensional variants.

Based on [BD77]:

Definition 3.1.1. *A De Morgan algebra over a set L is an algebra with:*

- *two binary operations \wedge and \vee called connections corresponding to taking minimum and maximum. These operations should be distributive: $\forall x, y, z \in L, x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$.*
- *a unary operator \neg called the involution such that the following laws hold:*

- double negation: $\forall x \in L, \neg\neg x = x$,
- De Morgan laws: $\forall x, y \in L, \neg(x \vee y) = \neg x \wedge \neg y, \neg(x \wedge y) = \neg x \vee \neg y$.
- bottom 0 and top element 1 satisfying $\forall x \in L, x \wedge 0 = 0, x \vee 1 = 1$. Also, $\neg 0 = 1$ and $\neg 1 = 0$.

Definition 3.1.2. Let \mathbb{A} be a set containing at most countably many elements. This set is called the set of names because the elements $x \in \mathbb{A}$ serve as names for dimensions of the cubes.

It is possible to consider a certain De Morgan algebra generated by dimension names. In this algebra, the individual variables describe the lines of a cube but the composite lattice elements that are also present in the algebra allow to define values on the interior of a cube.

Definition 3.1.3. Let $I \subset \mathbb{A}$ be a finite subset. The free De Morgan algebra generated by I is the smallest De Morgan algebra that contains all expressions with elements of I , the operators \vee, \wedge, \neg and the bounding elements 0, 1. It can also be denoted by $dM(I)$ or \mathbb{I} .

The involution of an element x in a free De Morgan algebra is sometimes denoted by $1 - x$ instead of $\neg x$. When viewing the dimension variable x as a line in a cube, the involution of x represents the inverted line.

Definition 3.1.4 (Category of cubes). The category \mathcal{C} contains as objects finite subsets $I \subset \mathbb{A}$, called (dimension) names. Morphisms in this category are maps into De Morgan algebras. More precisely, given two objects $I, J \in \mathcal{C}$, $f \in \text{Hom}_{\mathcal{C}}(J, I)$ if and only if $f : I \rightarrow dM(J)$ is a set map.

- If $f \in \text{Hom}_{\mathcal{C}}(J, I), g \in \text{Hom}_{\mathcal{C}}(K, J)$, the composition $g \circ f$, written as fg , is defined as the composition $\hat{g} \circ f$, where \hat{g} is the extension of g to the free De Morgan algebra $dM(J)$.
- The identity map in this category is just the identity.

From now on, until the rest of this chapter, the symbol \mathcal{C} will be used to denote the category defined in def. 3.1.4. \mathcal{C} can be interpreted as an abstract higher-dimensional cube on which we can apply operations described by a De Morgan algebra. As a base category it has much more objects and morphisms than the category that was used to construct directed and reflexive graphs, see example 2.5.4. As a consequence, types in a presheaf model over this category will be more complicated than in ??.

Now a few examples will be given of morphisms that are present in this category and will be used in the rest of the text.

Example 3.1.5. Let $I \subseteq J$. If $i \notin I$ then the morphism $s_i : I \cup \{i\} \rightarrow I$ is defined as the set map associated to the inclusion $I \subset dM(I \cup \{i\})$. Given $i \in I$ and $r \in dM(J)$, the morphism $(i/r) : I \rightarrow dM(J)$ maps the dimension name i on the element r of the free De Morgan algebra $dM(J)$ and leaves other names untouched. These maps are also called substitutions although they are not exactly the same as context substitutions.

- The face maps are compositions of substitutions (i/b) where $b \in \{0, 1\}$. These maps give the subfaces or edges of a higher-dimensional cube.
- The substitutions $(i/1 - i)$ can be interpreted as reflections along the i -axis or dimension of a cube. When we will model the intensional identity type with higher-dimensional cubes, the terms will correspond with edges on a cube or paths and this substitution can be used to model the reverse path or equality.
- The substitutions $(i/i \wedge j)$ extend a line parametrized by i to a square depending on j as well.

In presheaf models of type theory, the contexts, types and terms of the type theory are modelled with presheaves. In the cubical set model, the presheaves are taken over \mathcal{C} .

Definition 3.1.6. A cubical set Γ is a presheaf $\Gamma \in Psh(\mathcal{C})$.

Given two finite sets of dimension variables $I, J \in \mathcal{C}$ and a cubical set Γ , there is an element $\rho \in \Gamma(I)$. Using the assumption that Γ is a covariant functor, every substitution $f \in \text{Hom}_{\mathcal{C}}(I, J)$ gives a set map $\Gamma(f) : \Gamma(I) \rightarrow \Gamma(J)$. The application of $\Gamma(f)$ to the object $u \in \Gamma(I)$ is written on the right and the context Γ is implicit: $\rho f \in \Gamma(J)$.

Example 3.1.7. If $I = \{x, y\}$, we can interpret $u \in \Gamma(I)$ as a square with for example the edge $u(x/0)$ and corner $u(x/0)(y/0)$. In general, $u \in \Gamma(J)$ behaves like a $|J|$ -dimensional cube and the face maps give faces of this cube.

Example 3.1.8. Let R be a commutative ring with a unit element. Define for I a finite set of unknowns $\{x_1, \dots, x_n\}$, the set $R[I]$ of all polynomials in these unknowns. Then $R[-]$ can be considered as a functor and even a cubical set. For example if $p(x, y, z) = 1 + x^2y + z \in R[x, y, z]$, we can apply the face map $(y/0)$ to obtain $(1 + x^2y + z)(y/0) = 1 + z$.

Definition 3.1.9. The face lattice \mathbb{F} is the distributive lattice generated by the face maps (which are the substitutions of the form (i/b) , $i \in \mathbb{A}$, $b \in \{0, 1\}$). An element of the lattice $\varphi \in \mathbb{F}$ is called an extent. The top element in this lattice is denoted by $1_{\mathbb{F}}$ and the bottom element by $0_{\mathbb{F}}$ but the reference to \mathbb{F} is often omitted in literature. The face lattice also satisfies $(i/0) \wedge (i/1) = 0_{\mathbb{F}}$.

It is possible to define a cubical set using the face lattice. For a $I \in \mathcal{C}$, we set $\mathbb{F}(I)$ to be the subset of extents in \mathbb{F} that only contains dimension names $i \in I$. If $f \in \text{Hom}_{\mathcal{C}}(J, I)$, $\mathbb{F}(I) \rightarrow \mathbb{F}(J) : \varphi \mapsto f(\varphi)$ where f just substitutes the symbols in φ and applies the substitution rules on [Ort19], page 39. On the other hand, \mathbb{F} can also be seen as a type in the presheaf model over \mathcal{C} .

Remember from def. 2.5.5 that in such a presheaf model, contexts are presheaves and types are presheaves over a category with elements. The contexts in the presheaf model over the cube category \mathcal{C} are by def. 3.1.6 denoted by “cubical sets”.

The interpretation of the empty context $()$ in this presheaf model is by prop. 2.5.6 a presheaf denoted by 1 that maps all objects in the base category \mathcal{C} to the same set $\{\star\}$.

To define \mathbb{F} as a type in the empty context within the presheaf model, we need to describe how \mathbb{F} acts on the category of elements. By definition of the category with elements, we obtain that

$$\int_{\mathcal{C}} 1 = \{(I, \rho) \mid I \in \mathcal{C}, \rho \in 1I\} \cong \{(I, \star) \mid I \in \mathcal{C}\} \cong \mathcal{C}$$

where the last isomorphism is just a bijection that works by sending (I, \star) to the set $\mathbb{F}(I)$.

Definition 3.1.10. *The interval object \mathbb{I} represents the unit interval. It is modelled as a context in the presheaf model. Such a model is a functor $\mathcal{C} \rightarrow \mathbf{Set} : I \mapsto dM(I)$. $(i = 0)$ acts as a term $\mathbb{I} \vdash (i = 0) : \mathbb{F}$ on elements $\rho \in dM(I)$. This map is defined as the map setting all the occurrences of i in ρ to 0 and applying the rules of the distributive lattice.*

As it goes for every context, it can also be seen as a type in the empty context $() \vdash \mathbb{I}$. In this interpretation it is a functor $\int_{\mathcal{C}} \Gamma \rightarrow \mathbf{Set} : (I, \rho) \mapsto dM(I)$.

3.2 Manipulating contexts

It would now be possible to describe how traditional types that are defined in (univalent type theory) can be interpreted in this particular presheaf model. However, the face and interval objects form a very important part of cubical type theory because of their interaction with other contexts. To understand this interaction, it is necessary how types interact with contexts in general.

The presheaf model over \mathcal{C} also contains the context extensions from prop. 2.5.10: if $\Gamma \vdash A$ is a type, then the context extension is defined by

$$(\Gamma.A)(I) = \{(\rho, u) \mid \rho \in \Gamma(I), u \in A(I, \rho)\}.$$

Given $\Gamma \vdash A$ and an interpretation in the presheaf model, there is always a term q in

$$\mathbb{I}, A \vdash q : A.$$

This is used in cubical type theory in combination with the interval object to parameterise terms and types by dimension variables. For example if Γ is any context, there is the type $\Gamma \vdash \mathbb{I}$ and an extension $\Gamma.\mathbb{I}$ but also a term $\Gamma.\mathbb{I} \vdash q : \mathbb{I}$. $r(\Gamma, q : \mathbb{I})$ where q is as before but depends on the context Γ and type \mathbb{I} . In this situation and in literature q is replaced by a letter i, j, k and written in the context extension part. More precisely one writes $\Gamma, (i : \mathbb{I})$ instead of $\Gamma.\mathbb{I}$. If

In the presheaf model on cubical sets, there is another way to form new contexts using extents. But to introduce this formally, we need to pinpoint how extents are interpreted as terms.

Definition 3.2.1. *In general, the type $\Gamma \vdash \mathbb{F}$ is defined in the presheaf model by the contravariant functor*

$$\mathbb{F} : \{(I, \rho) \mid I \in \mathcal{C}, \rho \in \Gamma(I)\} \rightarrow \mathbf{Set} : (I, \rho) \mapsto \mathbb{F}(I)$$

where ρ is just ignored. Terms $\Gamma \vdash \varphi : \mathbb{F}$ are defined by a dependent function

$$\varphi : (I : \mathcal{C}) \rightarrow (\rho : \Gamma(I)) \rightarrow \mathbb{F}(I, \rho) : I \mapsto \rho \mapsto \varphi \rho$$

where I is implicit.

To calculate with φ , we need a precise interpretation in the presheaf and categories with families model. We define $0_{\mathbb{F}}, 1_{\mathbb{F}} \in \mathbb{F}$ to be terms $0, 1 : \mathbb{F}$ in the model by setting $0\rho = 0, 1\rho = 1$ for all $\rho \in \Gamma(I)$, $\Gamma \in \text{Psh}(\mathcal{C})$ and $I \in \mathcal{C}$.

For other extents, the definition or interpretation may depend on the context. When φ is an extent containing only variables $i, j, k \in \mathbb{A}$, there is a term $(i : \mathbb{I}), (j : \mathbb{I}), (k : \mathbb{I}) \vdash \varphi : \mathbb{F}$ also denoted $(i, j, k : \mathbb{I}) \vdash \varphi : \mathbb{F}$. It's interpretation in the category with families is as follows. Take an $I \in \mathcal{C}$ and $\rho \in (i, j, k : \mathbb{I})$. Then $\rho = ((\rho_i, \rho_j), \rho_k)$, where $\rho_i \in dM(I)$, $\rho_j \in \mathbb{F}(I, \rho_i)$ which means that ρ_j is an extent that only uses the variables in I . Further on, $\rho_k \in \mathbb{F}(I, (\rho_i, \rho_j))$. Now we can define $\varphi(I, \rho) = \varphi(I, (\rho_i, \rho_j, \rho_k))$ as the extent where we replace every variable i occuring in φ by the corresponding ρ_i and apply the rules on [Ort19], page 39.

Definition 3.2.2. Given a context Γ and term $\Gamma \vdash \varphi : \mathbb{F}$, we can define a new context (Γ, φ) , called the context restriction by extent φ . It is defined as

$$(\Gamma, \varphi)(I) = \{\rho \in \Gamma(I) \mid \varphi\rho = 1_{\mathbb{F}} \in \mathbb{F}(I)\}.$$

Example 3.2.3. Let $\Delta \equiv (\mathbb{I}, (i = 0))$. Then we can prove that $\Delta \cong ()$ as contexts.

Proof. We have to proof that there is a natural transformation between the two functors. First we have to prove that for all $I \in \mathcal{C}$, the sets $\Delta(I)$ and $()(I) = \{\star\}$ are isomorphic. In other words we have to prove that the set $\Delta(I)$ exactly contains one element. But by definition of the interval object and the context restriction $\Delta(I) = \{\rho \in dM(I) \mid (i = 0)\rho = 1\}$. But the only way that this could lead to 1 is when $\rho = 1$. So $\Delta(I)$ contains one element: $\rho = i \vee 1$. We still have to verify the commutativity properties of the natural transformation. \square

Similarly, it can be proven that the context $(i : \mathbb{I}, i = 0, i = 0)$ is isomorphic to the context $(i : \mathbb{I}, i = 0)$.

There is an inclusion of the restricted context in the original context using a context morphism $\iota_{\varphi} : (\Gamma, \varphi) \rightarrow \Gamma$. Intuitively, this restricted context can be used to define types and terms that are only defined the faces of the higher-dimensional cube Γ specified by the extent φ .

Example 3.2.4. A term $i : \mathbb{I}, j : \mathbb{I} \vdash a : A$ can be thought of as a square that is indexed by dimension variables i and j . The top and left sides of the square are described by the lattice formula $\varphi = (i = 0) \vee (j = 1)$. So if we want to restrict our attention to these sides, we can look at a in the restriction of the original context: $i, j : \mathbb{I}, \varphi \vdash a : A$. This is illustrated in fig. 3.1.

The operation of extending (actually restricting) a context by an extent, can be viewed as a typing rule:

$$\frac{\Gamma \vdash \varphi : \mathbb{F}}{\Gamma, \varphi \vdash}$$

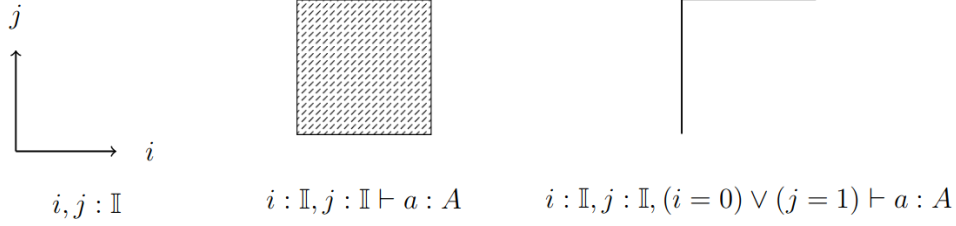


Figure 3.1: An illustration of the context restriction by $\varphi = (i = 0) \vee (j = 1)$ coming from [Ort19].

Definition 3.2.5. Let $\Gamma \vdash A$ and $\Gamma \vdash \varphi : \mathbb{F}$. A term u is called a partial element of extent φ if it is a term $(\Gamma, \varphi) \vdash u : A|_{\varphi}$.

There is also another definition of a partial element [Ort19]. This alternative definition let's the choice of u depend on $\rho \in \Gamma(I)$.

Definition 3.2.6. Let Γ be context and $I \in \mathcal{C}, i \notin I, \rho \in \Gamma(I \cup \{i\}), \varphi \in \mathbb{F}(I)$. A partial element u of $A(\rho)$ is a family $u_f \in A(\rho f)$ for every $f : J \rightarrow I \cup \{i\}$ such that $\mathbb{F}(s_i \circ f)(\varphi) = 1_{\mathbb{F}}$, with the property that for any $g : K \rightarrow J$ we have $u_{f \circ g} = A(g)(u_f)$. We call φ the extent of the partial element u .

For each $\Gamma \vdash u : A$ and extent φ , there is partial element of extent φ , or term $(\Gamma, \varphi) \vdash u|_{\varphi} : A|_{\varphi}$. When a term $(\Gamma, \varphi) \vdash v : A|_{\varphi}$ is induced by a restriction by an extent φ , or in other words $v = u|_{\varphi}$, we call v *extensible*. This can be intuitively seen as v being extendible to u by dropping the restriction by the extent φ . In type notation this relation between u and v is written as $\Gamma \vdash u : A[\varphi \mapsto v]$ expressing that φ extends the term v through φ . This may clarify the use of the word extent.

This specific notation for extents does not really state the presence of new types, contexts or terms in the model but can also be seen as a typing rule and written with natural deduction:

$$\frac{\Gamma \vdash v : A \quad \Gamma, \varphi \vdash v = u : A}{\Gamma \vdash u : A [\varphi \mapsto v]}$$

A generalization of extents in [Hub16a], page 95 is the following:

$$\frac{\Gamma \vdash a : A \quad \Gamma, \varphi_i \vdash a = u_i : A, \quad i = 1, \dots, k}{\Gamma \vdash a : A [\varphi_1 \mapsto u_1, \dots, \varphi_k \mapsto u_k]}$$

Definition 3.2.7. Assume that we have the following terms given:

- for each $1 \leq i \leq n$, there is a term $\Gamma, \varphi_i \vdash t_i : A$ that is only defined on the extent φ_i ,
- the extents cover the whole hypercube: $\Gamma \vdash \varphi_1 \vee \dots \vee \varphi_n = 1_{\mathbb{F}} : \mathbb{F}$
- the definitions are compatible $\Gamma, \varphi_i \vdash t_i = t_j : A, \forall 1 \leq i, j \leq n$,

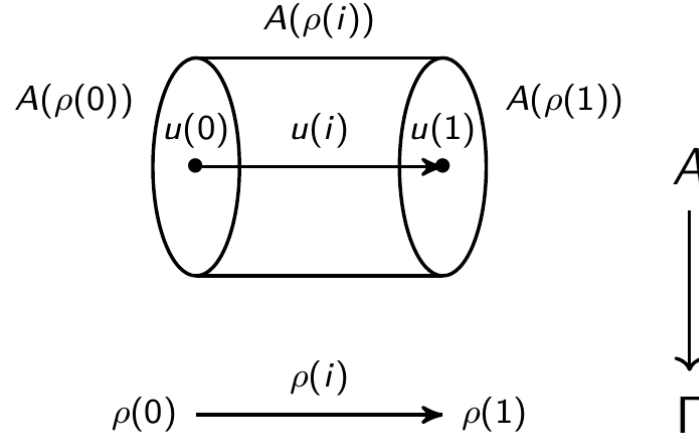


Figure 3.2: When $\Gamma \vdash A$ is a type in the presheaf model over cubical sets, it is described by a functor from the category with elements $\int_{\mathcal{C}} \Gamma$ to **Set**. This is noted as $A(\rho) \in \mathbf{Set}$ for some $|I|$ -dimensional hypercube $\rho \in \Gamma(I)$. At the same time the substitutions that apply to contexts, can be lifted to the level of types. If I is one dimensional, substitutions that map the dimension variables on 0 or 1 can be visualized as the sides of the cylinder. Picture taken from [Hub16b].

a system allows us to treat all the terms as one types and is denoted by $[\varphi_1 t_1, \dots, \varphi_n t_n] : A$.

Lemma 3.2.8. *Let $\Gamma \vdash \varphi : \mathbb{F}$ and $\Gamma \vdash A$. If $\varphi = \bigvee_i \varphi_i$ and $\Gamma \vdash \varphi_1 \vee \dots \vee \varphi_n = 1_{\mathbb{F}} : \mathbb{F}$, then the judgment that “ $A[\varphi \mapsto [\varphi_1 t_1, \dots, \varphi_n t_n]]$ is a type” in the category with families is equivalent with the judgment that “ $A[\varphi_1 \mapsto t_1, \dots, \varphi_n \mapsto t_n]$ is a type” in the category with families.*

Proof. Follows from the definition of generalized extents and the typing rule in [Hub16a], p. 95, figure 6.4 which states for any valid expression J :

$$\frac{\Gamma, \varphi_i \vdash J, 1 \leq i \leq n \quad \Gamma \vdash \varphi_1 \vee \dots \vee \varphi_n = 1_{\mathbb{F}} : \mathbb{F}}{\Gamma \vdash J}$$

□

For example the notation $[(i = 0) \vee (i = 1) \mapsto [(i = 0) u, (i = 1)v]]$ can be replaced by $[i = 0 \mapsto u, i = 1 \mapsto v]$. This is done very often in [Hub16a] and implementations [MV18].

3.3 Adding operations

We can now study the category theoretical interpretations of other terms and types in the presheaf model of cubical sets. In the presheaf model over the cube category, types get a intuitive description, see fig. 3.2.

However, we would first of all like to say something about univalent type theory. The interesting properties of univalent type theory and the consequences of the univalence axiom all follow from the intensional identity type and its typing rules. The intensional

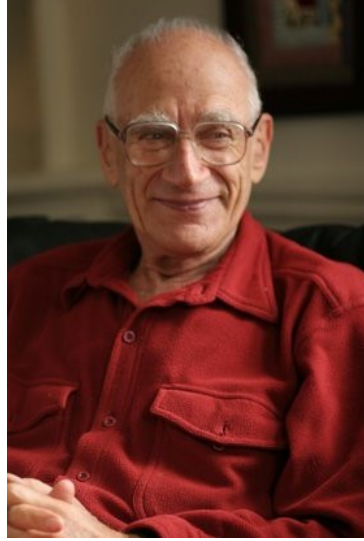


Figure 3.3: Daniel Kan (1927-2013), here seen in his home in Massachusetts, was a Jewish category and homotopy theorist. He defined a kind of Kan extension property for the first time in the context of algebraic topology for cubical sets, see his book [Kan55] saying how boxes can be filled. Although [CHS19] mentions a predecessor to this property in [Eil39]. This Kan extension property later involved into different flavors such as a version for simplicial sets stating that every “horn” can be filled.

identity type of univalent type theory can be modelled in cubical sets but it is special in the following sense. Given a type A and terms $a, b, c : A$, the interpretation of the intensional identity type $a =_A b$ and $b =_A c$ in cubical sets is only transitive if A satisfies a certain *Kan extension property* (see fig. 3.3). This property of types A tells us how the terms of A can be extended from a restricted context cube to a full cube. It is made precise with a composition and filling operation on the types and terms in the presheaf model with cubical sets.

Adding *operations* to the the model is done by adding axioms stating the existence of certain structures, sets, functors or elements, satisfying properties that are related to the properties of the type system that is modelled. Operations will be denoted with another font type. To arrive at a complete model of univalent type theory, including the univalence axiom, also a glue operation has to be added (see section 3.3.4).

3.3.1 The composition operation

The first and most important operation is the composition operation [Ort19].

Definition 3.3.1. A composition structure for $\Gamma \vdash A$ is an operation, comp , that depends on the following elements:

- $I \in \mathcal{C}, i \notin I, \rho \in \Gamma(I \cup \{i\}), \varphi \in \mathbb{F}(I)$, where ρ can be seen as an $(I+1)$ -dimensional hypercube.
- u a partial element of $A(\rho)$ of extent φ and $a_0 \in A(\rho(i/0))$ such that for all $f : J \rightarrow I$, $A(f)(a_0) = u_{(i/0) \circ f}$ and $\mathbb{F}(f)(\varphi) = 1_{\mathbb{F}}$. Here u can be viewed as a set

of values for A on sides φ of ρ that takes values a_0 on the bottom of ρ described by $i = 0$. The value a_0 has to agree with the values u on other sides.

We have an element $\text{comp}(I, i, \rho, \varphi, u, a_0) \in A(\rho(i/1))$. This element is uniform in the sense that for any $f : J \rightarrow I$ and $j \notin J$ we have

- $\text{comp}(I, i, \rho, \varphi, u, a_0) f = \text{comp}(J, j, \rho(f \cup (i/j)), \varphi f, u(f \cup (i/j)), a_0 f)$ where $f \cup (i/j) : J \cup \{j\} \rightarrow I \cup \{i\}$ is just the extension of f with the substitution (i/j) and $u(f \cup (i/j))$ is the partial element defined for $g : K \rightarrow J$ as $u_{(f \circ g \cup (i/j))}$.
- $\text{comp}(I, i, \rho, 1_{\mathbb{F}}, u, a_0) = u_{(i/1)}$. In other words, the values of A on top of the hypercube ρ can be extended to values on top of the faces of the hypercube ρ for which $i = 1$. In other words, it is possible to close the “lid” of the box ρ along the side $i = 1$.

Again we can write this operation as a typing rule telling us the existence of the category theoretical interpretation in the model of a specific term with a specific type. The notation $\text{comp}(I, i, \rho, \varphi, u, a_0)$ is written with a shorter syntax and with application to ρ implicit as $\text{comp}^i A [\varphi \mapsto u] a_0$.

$$\frac{\Gamma \vdash \varphi : \mathbb{F} \quad \Gamma, i : \mathbb{I} \vdash A \quad \Gamma, \varphi, i : \mathbb{I} \vdash u : A \quad \Gamma \vdash a_0 : A(i/0)[\varphi \mapsto u(i/0)]}{\Gamma \vdash \text{comp}^i A [\varphi \mapsto u] a_0 : A(i/1)[\varphi \mapsto u(i/1)]}$$

If $\sigma : \Delta \rightarrow \Gamma$ is a context substitution, then this morphism acts on terms as follows:

$$\frac{(\text{comp}^i A [\varphi \mapsto u] a_0) \sigma}{\text{comp}^j A (\sigma, i/j) [\varphi \sigma \mapsto u(\sigma, i/j)] a_0 \sigma}$$

where j does not appear as a dimension variable in a syntactical representation of Δ yet. For types the morphism acts analogously.

The existence of a composition structure cannot be taken for granted. Its existence depends on the choice of context and type.

Definition 3.3.2. A Kan type or fibrant type A is a type for which there exists a composition structure comp_A .

3.3.2 The path type

The identity type in cubical type theory is sometimes called path type. This is because cubical type theory is modelled with presheaf models and the standard model of the identity type is defined as a functor Id in the following way.

If $\Gamma \vdash A$, $\Gamma \vdash a : A$ and $\Gamma \vdash b : A$, then

$$(Id_A(a, b)) \rho = \{\star \mid a\rho = b\rho\}$$

However, the type modelled by the functor Id does not satisfy the properties of the intensional identity type in univalent type theory. It does for example only admit at most one term and satisfies the uniqueness of identity proofs principle (and also

Work out.

ax. 2.5.14). This means it is incompatible with the univalence axiom and insufficient to model univalent type theory.

To solve these problems In cubical type theory and other presheaf models, another interpretation is given to the identity type. This interpretation is done with a functor that models the intensional identity type between terms $a, b : A$ correctly (up to the elimination rule) is called $\text{Path}_A(a, b)$.

Definition 3.3.3. *Given terms $\Gamma \vdash a, b : A$, the path type $\text{Path}_A(a, b)$ is defined as follows:*

- *Given $\rho \in \Gamma(I)$, $\text{Path}_A(a, b)(\rho)$ is the set of equivalence classes generated by pairs (i, w) with $i \notin I$ and $w \in A(\rho s_i)$ such that $w(i/0) = a(\rho)$ and $w(i/1) = b(\rho)$, where we identify (i, w) with (i', w') iff $w' = w(i/j)$.*
- *The action of $\text{Path}_A(a, b)$ on a morphism $f : (J, \rho, f) \rightarrow (I, \rho)$ is given by $(i, w)f \equiv (j, w(f \cap (i/j)))$ for $j \notin J$.*

The equivalence class of (i, w) is also denoted by $\langle i \rangle w$ where w implicitly depends on a choice of ρ . Formally, using typing rules and natural deduction:

$$\frac{\Gamma \vdash A \quad \Gamma, i : \mathbb{I} \vdash t : A \quad \Gamma, i : \mathbb{I} \vdash t(i/0) = a : A \quad \Gamma, i : \mathbb{I} \vdash t(i/1) = b : A}{\Gamma \vdash \langle i \rangle t : \text{Path}_A(a, b)}$$

$$\frac{\Gamma \vdash t : \text{Path}_A(u_0, u_1)}{\Gamma \vdash t0 = u_0 : A}$$

$$\frac{\Gamma \vdash t : \text{Path}_A(u_0, u_1)}{\Gamma \vdash t1 = u_1 : A}$$

To obtain a model for intensional identity type that also satisfies the elimination rule of the intensional identity type (see), the identity type of Swan [Swa14] has to be used instead of the path type in def. 3.3.3. The identity type of Swan does exactly have the elimination rule of . However, the difference between the path type in def. 3.3.3 and the identity type of Swan is only minor (up to paths) and will not be covered thoroughly in this text. Both Swan's identity type and the path identity type in def. 3.3.3 are logically equivalent, meaning that they can be substituted by each other. See [Hub16a], p. 114 for a comparison of both types and the more advanced text [Swa18] about a general framework for choosing identity types. In this text, the notation $a =_A b$ or simply $a = b$ will be used for both Swan's identity type (constructed by Swan) and the path identity type between terms $a, b : A$ defined in this section.

The standard term of the path type $\text{Path}_A(a, a)$ between a term $a : A$ and itself is the reflexive term $\text{refl}_A(a)(I, \rho) \equiv (i, a(\rho s_i))$ for some $i \notin I$. It is in [Hub16a], page 90 also denoted by 1_a .

This definition applies to any type, but to have that the path type is Kan type and satisfies all the properties of the intensional identity type in univalent type theory, we need that the original type to which we apply the path type is Kan. If the original type was not Kan, the path would for example not necessarily be transitive.

$$\begin{array}{ccc}
a & \dashrightarrow & c \\
\uparrow \text{refl} & & \uparrow q \ j \\
a & \xrightarrow{p \ i} & b
\end{array}$$

Figure 3.4: Transitivity can be modelled in cubical type theory by a composition of paths along the edges of a square. Diagram inspired by [CCHM16], section 4.3.

Example 3.3.4. *The transitivity property of equality, in this case the path type, states that if $\Gamma \vdash a, b, c : A$, $p : \text{Path}_A(a, b)$ and $q : \text{Path}_A(b, c)$, there is a term $r : \text{Path}_A(a, c)$. If A is Kan, then Path is transitive.*

Proof. Using the reflexive term, we obtain fig. 3.4.

If A is a Kan type, by definition, it has a composition operation. The syntactical object $u = [i = 0 \mapsto a, i = 1 \mapsto qj]$ can be viewed as a term that is only defined on three sides:

- The extent $\varphi = (i = 0) \vee (i = 1)$ defines two sides of the hypercube. On these sides the values of u vary along a second dimension j . This is written as $\Gamma, (i = 0) \vee (i = 1), j : \mathbb{I} \vdash u$.
- On the side $j = 0$ the values of u coincide with the values of the path p . This means that for fixed i , in type notation $\Gamma \vdash (p \ i) : A(j/0)[\varphi \mapsto u(j/0)]$

For fixed i , we take $a_0 = p \ i$. The composition operation now gives us a term that extends u along the remaining side $j = 1$:

$$\text{comp}^j A [(i/0 \mapsto a, (i/1) \mapsto q \ j)](p \ i)$$

which can be rewritten using lemma 3.2.8 as

$$\text{comp}^j A [(i/0 \mapsto a, (i/1) \mapsto q \ j)](p \ i) \equiv z.$$

The term z is of type $A(j/1)[i = 0 \mapsto a, i = 1 \mapsto c]$.

If we let this term depend on i , we obtain with the introduction rule of a path, a term denoted by $\langle i \rangle z$ that has type $\text{Path}_A(a, c)$. To verify this we need to check that the endpoints of this path $\langle i \rangle z$ are indeed a and c .

The endpoints are obtained by applying the context morphisms

$$(i/0), (i/1) : (\Gamma.(j : \mathbb{I})) \rightarrow (\Gamma.(j : \mathbb{I}).(i : \mathbb{I}))$$

to the generalized extent judgement

$$(\Gamma.(j : \mathbb{I}).(i : \mathbb{I})) \vdash z : A(j/1)[i = 0 \mapsto a, i = 1 \mapsto c].$$

The application of a context morphism σ is just the application of σ to each judgment contained in the generalized extent, which are by definition (omitting $(\Gamma.(j : \mathbb{I}).(i : \mathbb{I}))$):

- $() \vdash z : A(j/1)$

$$\begin{array}{ccc}
 a & \xrightarrow{p} & b \\
 p(i/0) \uparrow & & \uparrow p(i/j) \\
 a & \xrightarrow{p(j/0)} & a
 \end{array}$$

Figure 3.5: Take two terms $a, b : A$ and a path $p : \text{Path}_A(a, b)$, then the term $p(i/i \wedge j)$ is a two-dimensional path that takes two arguments i, j and returns values based on the interpretation of $i \wedge j$ as $\min(i, j)$.

- $(i = 0) \vdash z = a : A(j/1)$
- $(i = 1) \vdash z = c : A(j/1)$

Now we have proven that z is the right path, making the transitivity property hold for Path . □

If A is Kan, it can be proven that the path type $\text{Path}_A(a, b)$ not only satisfies transitivity but also all the typing rules of the intensional identity type (as mentioned in def. 1.2.5 or listed on page 35 in [Ort19]) except the elimination rule. However, it is not known how much the Kan property necessary for this property of the identity type, and whether the composition operation can be weakened.

In [CM19], the composition operation is weakened to a new kind of composition operation, a weak composition. This allows to develop a type theory in which the mathematical properties of the different cubical type theories can be studied simultaneously, see for example the alternative theory described in section 3.6.2.

3.3.3 The filling operation

The composition operation of a Kan type extends the definition of a partial term from one that defined only on one endpoint of a dimensional variable to one that is also defined on the other endpoint of the dimension variable. It however only gives us the lid of the box as described in example 3.3.4 which is actually the example on [CHM18], page 10.

Complementary to the composition operation, there is a filling operation that not only “fills” the endpoints but also covers the interior of partial terms that are defined on hypercubes by introducing a new variable dependency in the context. This operation is used in [CHM18] for the definition of higher inductive types. The filling operation makes use of the lattice structure that is defined in def. 3.1.1.

Definition 3.3.5. *Let A be a Kan type. The Kan filling operation on A is an operation that returns a term. More specifically, we set $\text{fill}^i A \ [\varphi \mapsto a_0] \ a_0$ to be the term*

$$\text{comp}^j A(i/i \wedge j) \ [\varphi \mapsto u(i/i \wedge j), (i = 0) \mapsto a_0] \ a_0 : A$$

in context $\Gamma.(i, j : \mathbb{I})$.

Using the definition of extents, we can decompose the filling term v into judgments (omitting $\Gamma, i, j : \mathbb{I}$):

- The judgements $() \vdash v : A(i/i \wedge j)$ and $\varphi \vdash v = u(i/i \wedge j) : A(i/i \wedge j)$. This means that the term v is defined on the interior $A(i/i \wedge j)$. If we apply (j/i) we obtain the usual composition term.
- There is a judgment $(i = 0) \vdash v = a_0 : A(i/i \wedge j)$ that is equivalent with

$$() \vdash v(i/0) = a_0 : A(i/0).$$

Now $a_0\rho$ is on the left side of the hypercube $\rho \in \Gamma(\{i, j\})$.

The filling is the composition plus something else, also called the interior. But the interior in this context is not the interior in a topological sense because we do not have any topology. In the setting of cubical sets we can only compute with a finite amount of (subsets of) corners of standardized hypercubes.

The existence of Kan types, types with a composition operation is not so straightforward. For each type that is used in a type system and has a valid interpretation in the category with families model, the composition operation has to be defined. Examples of such composition structures for natural numbers and path types can be found in [Hub16a], section 6.4.5, p. 97-98. Once these composition structures are given for all basic types in intensional type theory, including paths, the category with families model is strong enough to model univalent type theory. However, the univalence axiom is an axiom about universes of intensional type theory and universes are also types. So we need a composition structure for universes. This composition structure is defined in [Hub16a] with the help of a `Glue` type which will be introduced in section 3.3.4.

3.3.4 The `Glue` type

The `Glue` type has typing rules that manipulate the types or types in def. 1.3.1 (originally from the book [VAC⁺13]). Now that the basic types for cubical type theory have been modeled by their semantics in cubical sets, these concepts have a cubical type-theoretical analogue. The definition of a term of contractibility in intensional type theory is

$$\text{isContr } A = (x : A) \times ((y : A) \rightarrow \text{Path}_A(x, y))$$

and is a valid expression in cubical type theory because sum types and path types have a valid interpretation in cubical sets.

Similarly, a function $f : T \rightarrow A$ is an *equivalence* if and only if the following type is inhabited

$$\text{isEquiv } T \ A \ f = (y : A) \rightarrow \text{isContr}((x : T) \times \text{Path}_A(y, f \ x))$$

and an equivalence in general is a term of the type

$$\text{Equiv } T \ A = (f : T \rightarrow A) \times \text{isEquiv } T \ A \ f.$$

There is a semantic interpretation of the `Glue` type in terms of cubical sets given in [Hub16a] but this will be skipped because it is quite complicated and does not add much intuition for the value of this type in the context of the proof of univalence. The text [Ort19], chapter 5-6, tries to simplify the semantics of the `Glue` type using *topoi*. The introduction of the `Glue` type is given in [Ort19], page 30.

Every type comes with typing rules. Assume that we are given a (total, not necessarily partial) type $\Gamma \vdash B$ and another type that is partial $\Gamma, \varphi \vdash A$ for some extent φ . If the partial type is equivalent to the total type $\Gamma, \varphi \vdash f : \text{Equiv } A \ B$ as far as it is defined, there is a type $\Gamma \vdash \text{Glue } [\varphi \mapsto (A, f)] \ B$ that satisfies certain elimination rules.

The formal formation rule of the Glue type tells that give an element of the face lattice \mathbb{F} expressing the sides of a square on which f is defined as an equivalence $A \rightarrow B$, there is a bigger type that “glues the sides together”:

$$\frac{\Gamma \vdash \varphi : \mathbb{F} \quad \Gamma \vdash B \quad \Gamma, \varphi \vdash A \quad \Gamma, \varphi \vdash f : \text{Equiv } A \ B}{\Gamma \vdash \text{Glue } [\varphi \mapsto (A, f)] \ B}$$

By one of the eliminating typing rules as mentioned on [Hub16a], the Glue type extends the partial type A . More precisely, it says that for $\varphi = 1_{\mathbb{F}}$, we have $\Gamma \vdash [1_{\mathbb{F}} \mapsto (T, f)] \ B = A$. Terms of the Glue type come into two forms.

- If it known whether B has a particular partial type, there is a typing rule giving us a glue term. More precisely, if t is a term of A , f is the equivalence and $\Gamma \vdash b : B \ [\varphi \mapsto f \ a]$, then there is a term denoted by $\text{glue } [\varphi \mapsto a] \ b$ of the glue type.
- The unglue term functions as an inverse of this term, a kind of eliminator typing rule. For example, given a Glue term $c : \text{Glue } [\varphi \mapsto (A, f)] \ B$, the term $\text{unglue } c$ is of type $B \ [\varphi \mapsto f \ c]$.

Example 3.3.6. Take $\varphi = (i = 0) \vee (i = 1)$. Then the type A is only defined for $i = 0$ or $i = 1$, so there are two sides of the hypercube on top of which it is defined. The values on top of these sides are called A_0 and A_1 . The values of B are computed with substitutions. The partial equivalence also has values f_0 and f_1 . If we call $\mathbb{G} = \text{Glue } [i = 0 \mapsto (B(i/0), f_0), i = 1 \mapsto (B(i/1), f_1)]$, this gives a diagram:

$$\begin{array}{ccc} A_0 & \xrightarrow{\quad \mathbb{G} \quad} & A_1 \\ \downarrow f_0 & & \downarrow f_1 \\ B(i/0) & \longrightarrow & B(i/1) \end{array}$$

get example on Hub16a, p. 101

As it goes for all the types that are introduced in cubical type theory, it is also necessary to introduce a composition operation for the Glue type. This is done on [Hub16a], section 6.6.2 and makes the Glue type into a Kan type.

The only remaining step to do is to prove that the universe type is also Kan, which is done in [Hub16a], section 6.7.2. Once the composition and glueing operations are defined for the universe types, they are defined for all types. As explained in example 3.3.4, this implies that the path type over universes also satisfies transitivity. This means the model described in this section supports all types from intensional type theory, the Glue -type and a composition structure:

Definition 3.3.7. The cubical type theory (as defined in [Hub16a]) is an extension of intensional type theory with the following properties:

- All types from intensional type theory, including natural numbers, dependent sum and product types, are modeled with a presheaf model over the cube category *def.* 3.1.4 by defining them as functors. This is done in the introduction of [Hub16a], p. 28-30.
- The intensional identity type is modeled by modifying the path type of which the semantics and typing rules are given in section 3.3.2.
- There are the following additional types that are not in intensional type theory or homotopy type theory:
 - The objects \mathbb{F} and \mathbb{I} introduced in *defs.* 3.1.9 and 3.1.10 and the context restriction *def.* 3.2.2.
 - The $\mathbb{G}\text{Lue}$ type from section 3.3.4 allows the construction of a proof of univalence, see section 3.4.
- Every type has a composition, satisfying the typing rules (and semantics) in *def.* 3.3.1, allowing the intensional identity type to be transitive and cubical type theory be a proper extension of intensional type theory.

The theory in *def.* 3.3.7 (in which the univalence axiom is provable) satisfies *canonicity of numerals* (see *def.* 3.3.8 which was proven in [Hub16a], published in [Hub17] and further refined in [CHS19]. This property means that expressions involving terms of the type of natural numbers should be (path) equal to an expression that uses only a finite amount of constructors. It is defined formally for the special case of natural numbers on [Hub16a], p. 125:

Definition 3.3.8. *given a context I of the form $i_1, \dots, i_k : \mathbb{I}$ and a derivation of $i \vdash u : \mathbb{N}$ (where \mathbb{N} is an inductive type), there is a unique $n \in \mathbb{N}$ (in a metatheory) with $I \vdash u = \text{suc}^n \text{zero} : \mathbb{N}$ for some $n \in \mathbb{N}$. This n can more-over be effectively calculated.*

When this property holds, at least every basic type (the numerals and similar types) has some standard form that can be computed but in practice computing this standard form does not always work, see section 3.5.3. This property may fail to hold when axioms such as function extensionality or univalence are added the theory. Cubical type theory is an example of a theory in which both canonicity and univalence holds. There is also another kind of canonicity in literature: “canonicity of booleans” [AHH18]. But because numeral values give rise to boolean expressions and the other way around, canonicity of booleans is equivalent with canonicity of numerals.

According to [Bru18] the proof of canonicity solved a conjecture by Voevodsky, originally formulated for homotopy type theory:

Given a term $t : \mathbb{N}$ constructed using the univalence axiom, we can construct two terms $u : \mathbb{N}$ and $p : t =_{\mathbb{N}} u$ such that u does not involve the univalence axiom.

3.4 Proof of univalence

Cubical type theory was invented to give a “computational interpretation of the univalence axiom” which would allow to compute with expressions in univalent type theory. The canonicity property proven in [Bru18] predicted that many computations should terminate in a finite amount of steps. However, this property was proven for def. 3.3.7 and not yet univalent type theory. The remaining problem to solve is proving that def. 3.3.7 is indeed a model of the univalence axiom. In this section, an overview will be given of historical attempts to this proof and one of the most recent versions of the proof.

3.4.1 Statement of theorem

Remember from ax. 1.3.2 that the *univalence* theorem in cubical type theory (which is an axiom in univalent type theory) states:

Theorem 3.4.1. *Given two types $A, B : \mathcal{U}$ of the same universe, there is an equivalence between the types $A = B$ and $A \simeq B$.*

Because it can be proven in certain models such as [KL12] and cubical type theory as in def. 3.3.7, the axiom can also be called a theorem within a particular model. In this section, the proof of the theorem in [MV18] will be discussed.

3.4.2 History of the proof

The first model in which this theorem became provable was simplicial sets [KL12]. The proof in simplicial sets used the ideas such as the Glue type that were reintroduced in later proofs in cubical sets, see [Hub16a], p. 103-104 and p. 122-123 or [BCH18] for an approach without lattice structure (mentioned in section 3.6.1).

The proof in [Hub16a] was the first one that was formally verified in a type theory, more specifically cubical type theory. It has since been rewritten in [Wei16]. The formalization has been continuously worked on since [CCHM15] and [CCHM16]. The most recent version can be found in the Agda library Agda-Cubical [MV18], `CubicalFoundations.Univalence`. Recently, this library code was surveyed in [Mö18a] but revisions of the code may have made this explanation less useful. In this section the most recent revision of the proof will be discussed.

is

3.4.3 Contractibility of equivalence singletons

Theorem 3.4.2 (`unglueIsEquiv`). *Let $\Gamma, \varphi \vdash f : T \rightarrow A$ be a partial equivalence with extent φ . The function*

$$\text{unglue} : \text{Glue } [\varphi \mapsto (T, f)] \ A \rightarrow A$$

is an equivalence.

Another proof is [Hub16a], p. 103, Theorem 6.7.2 and uses two extra lemmas to construct to proof the equivalence.

Proof. As mentioned in def. 1.3.1 and ax. 1.3.2, an equivalence is a function such that its fibers (pre-images) are contractible. So, to prove that `unglue` is an equivalence amount to proving that for any b in the co-domain A , the fiber `fiber unglue b` is contractible. In other words, we need to give proofs of the following:

- A point $x : \text{fiber unglue } b$.
- A proof that given $y : \text{fiber unglue } b$, $x = y$.

The point x is constructed by constructing the `glue` term on top of `hcomp u b : A` where u is a partial eterm. It has to be proven that this term is indeed in the fiber. For the second part, an arbitrary element of the fiber y is taken. Now a path is constructed that is point-wise defined with a partial term u' and `glue` terms. The end points of this path are indeed x and y because u' extends u . There also have to be proofs that the terms on the path are in the fiber (or pre-image of b). These proofs are built with an application of `hfill`, a formalization of section 3.3.3. \square

Now `unglueEquiv` expresses that any partial family of equivalences can be extended to a total one.

The following theorem is an intermediate statement of the univalence axiom that is more directly provable than the traditional statement. It will function as a lemma for proving the more traditional version of the univalence axiom.

Theorem 3.4.3 (`EquivContr`).

$$\forall A : \mathcal{U}, \text{isContr} \left(\sum (T : \mathcal{U}) (T \simeq A) \right).$$

The proof in [Hub16a], p. 104, Corollary 6.7.3 uses an extra lemma.

Proof. We need to give a term $x : \sum T : \mathcal{U} (T \simeq A)$ and a proof that there is an equality between x and any other term y of the same type. Take the trivial equivalence $(A, \text{idEquiv})$ for x . It remains to prove that every other term is equal to this one. So take another term

$$y : \Sigma(\text{Set}\ell)(\lambda T \rightarrow T \simeq A)$$

then it can be shown with `unglueIsEquiv` that er is a path. \square

According to [Hub16a], [VAC⁺13], Theorem 5.8.4, it should follow that the equivalence type forms an identity system. The statement of the univalence theorem is however not immediate.

The path type in cubical type theory can be slightly modified (see section 3.3.2) such that it satisfies the J-eliminator, see section 3.3.2. Equivalences also satisfy an eliminator, called *equivalence induction*.

Theorem 3.4.4 (`EquivJ`). *Given a property $P : (A B : \mathcal{U}) \rightarrow (e : B \simeq A) \rightarrow \mathcal{U}$, a proof of the base step $r : (A : \mathcal{U}) \rightarrow P A A (\text{idEquiv } A)$, then we can produce a proof of the property for any $A B : \mathcal{U}$:*

$$P A B e$$

.

Proof. For the proof, some lemmas are needed:

- The fact that $\text{idIsEquiv} : \forall A \in \mathcal{U}, \text{isEquiv } (\lambda x \mapsto x)$ gives a proof of $\text{idEquiv} : \forall A : \mathcal{U}, A \simeq A$ which can be seen as reflexivity for equivalences.
- Applying theorem 3.4.3 to some type B , we get a proof of contractibility of a sum type depending on B . More precisely, there is a $(C, C \simeq B)$ with $C : \mathcal{U}$ such that every other C' gives an equality $(C, C \simeq B) = (C', C' \simeq B)$. Contractibility is very similar to being an h-prop. The theorem $\text{isContr} \rightarrow \text{isProp}$ which is stated in `CubicalFoundations.HLevels` converts proofs of contractibility into proofs of being an h-prop using the composition operation. It shows that any terms of this particular sum type over B are connected by a path or equality.
- Assuming $e : A \simeq B$, the above theorem is applied to connect the tuples $(B, \text{idEquiv } B)$ and (A, e) by an equality in the proof of theorem `contrSingleEquiv`.
- Transport is given by

$$\text{subst} : (B : A \rightarrow \mathcal{U}) (p : x = y) \rightarrow B x \rightarrow B y.$$

This theorem is part of a few theorems that are all formalizations of univalent type theoretical transport in defined with the composition operation in `Cubical.Core.Prelude`.

Eventually the J-combinator for equivalences is defined by transporting the base case r over the equality returned by `contrSingleEquiv e`.

□

This construction of the map ua is based on the one defined in [Hub16a], Example 6.6.2, p. 101.

Definition 3.4.5. $\text{ua} : \forall A B : \mathcal{U}, A \simeq B \rightarrow A = B$ is defined with the `Glue` type. The definition can be illustrated with a diagram:

$$\begin{array}{ccc} A & \xrightarrow{\text{ua}} & B \\ \downarrow e & & \downarrow \text{idEquiv } B \\ B & \longrightarrow & B \end{array}$$

The implementation of the definition takes a partial equivalence e and type A , a dimension variable or parameter i and maps it on `Glue` $[(i = 0) \mapsto (A, e), (i = 1) \mapsto (B, \text{idEquiv } B)] B$. This is a path with endpoints A and B .

$\text{uaIdEquiv} : \forall A, \text{ua} (\text{idEquiv } A) = \text{refl}$, where `refl` is the reflexive path or equality. It tells us that the reflexive equivalence is mapped onto the reflexive identity by ua .

Definition 3.4.6. An isomorphism (or homotopy equivalence, quasi-inverse [VAC⁺ 13]) is defined in `CubicalFoundations.Isomorphism` as a function $f : B \rightarrow A$ with a pseudo-inverse $g : B \rightarrow A$ such that $f \circ g$ and $g \circ f$ are both homotopic to the identity map.

The theorem `isoToIsEquiv` is a proof that any isomorphism $f : A \rightarrow B$ is an equivalence.

Example 3.4.7. Let $A = \text{Bool}$ and $B = \text{Fin } 2$. We can prove that $A = B$. This is done using any bijection between A and B which is also an isomorphism. Such an isomorphism is also an equivalence and with the `Glue` type in ua we obtain a path of type $A = B$.

3.4.4 Conclusion of the proof

Using the previous lemmas and results, the statement and proof of the univalence theorem is now much easier.

Theorem 3.4.8 (`Univalence.thm`). *Given a function*

$$\text{au} : A = B \rightarrow A \simeq B$$

and a proof that `au` maps the `refl` : $A = A$ identity type constructor onto the constant equivalence $A \rightarrow A$, the theorem `Univalence.thm` states that `au` is an equivalence.

Proof. First it is shown that the given map `au` has to be an isomorphism with as inverse the previously constructed map `ua`. An isomorphism is the homotopy type theoretic variant of a homotopy equivalence or more precisely, a bi-invertable map as in chapter 4 of [VAC⁺13]. To prove that `au` is a isomorphism, we need to prove the left-inverse and right-inverse properties (up to a path). Both proofs use that `compPath` is just the transitivity property of “=” and congruence.

- The left-inverse is proven with equivalence induction `EquivJ` on the property that `au` is a left-inverse for equivalences `e`: `au (ua e) = e`.
- The right-inverse is proven with equality induction `J` on the property that `ua` is a right-inverse for paths `p`: `ua (au p) = p`.

Although an isomorphism is not well-behaved (it is for example not a proposition), it contains more information than equivalence and from an isomorphism, an equivalences can be extracted. Applying `isoToIsEquiv` gives that `ua` : $A = B \rightarrow A \simeq B$ is an equivalence or `isEquiv ua`. \square

The theorem of univalence is then a simple consequence of an application of `Univalence.thm`

Theorem 3.4.9 (`univalence`). *Given any $A, B : \mathcal{U}$, there is an equivalence $(A \equiv B) \simeq (A \simeq B)$.*

Proof. The function `lineToEquiv` : $A = B \rightarrow A \equiv B$ that maps paths directly on an equivalence serves as a candidate for the inverse `au` in `Univalence.thm`. The necessary condition that is still needed to apply `Univalence.thm` is the fact that `lineToEquiv` maps the reflexive equality onto the trivial equivalence. This follows by proving that the underlying maps of the image under `lineToEquiv`, `pathToEquiv refl` and `idEquiv A` are equal with transport and using the fact that the property of being an equivalence is a proposition, which means there can only be one such proof, such that the the equivalences also have to be equal and `pathToEquiv refl = idEquiv A`. The proof is formalized in `pathToEquivRef1`. Applying `Univalence.thm` to the map and this fact, gives us a proof that `pathToEquiv` is an equivalence. Assembling this into an equivalence gives a term of the equivalence type $(A \equiv B) \simeq (A \simeq B)$, which is a proof of univalence. \square

3.4.5 Univalence with topoi

The *language of toposes* can be used to discover more fundamental models of cubical type theory and univalence [Ort19]. A topos is a category that behaves like the category of sets. Examples of topoi are the category of sets and presheaves.

First, the cubical type theory as presented in [Hub16a] and this text is translated in the language of toposes [Ort19], section 5.3. Some definitions become much simpler such as the De Morgan algebra used for the base category and the face lattice that were presented in def. 3.1.9 become [Ort19], figure 5.4. This results in an alternative proof of univalence that may be more insightful [Ort19], section 5.5. An overview of related foundational work can be found in [Pit18] and [LOPS18].

3.5 Applications of the proof of univalence

In section 3.4, it was shown that the univalence axiom holds in cubical type theory. The theorem of univalence allows us to use all the concepts from univalent type theory, also called homotopy type theory, in cubical type theory. Cubical type theory was implemented in `Cubical.Core.HoTT-UF` of [MV18].

In combination with the standard library of Agda [DDG⁺19], it is possible to produce concrete examples in univalent type theory of reasoning with univalence.

3.5.1 Isomorphism invariant algebra

In the theories of groups and other algebraic structures, proofs are preferably done for only one structure in an isomorphism class. The proofs for other structures in the same class are not considered valuable because they do not add information. With the univalence axiom, it is possible to treat these different proofs as equal.

This can be shown in the context of monoids for the following two concrete examples:

$$s_1 \equiv (\mathbb{N}, (m, n) \mapsto m + n, 0)$$

and

$$s_2 \equiv (\mathbb{N}_0, (m, n) \mapsto m + n - 1, 1)$$

from [CD13] and [Dan12].

The goal of this example is to create a term of the path type between s_1 and s_2 in `Mon` and transport proven properties of s_1 from the type of properties on s_1 to the one on s_2 without repeating the proof for s_2 . For example, if we have a proof of the fact that s_1 is commutative, we also have a proof of the fact that s_2 is commutative. This shows one application for the proof of univalence: the possibility of doing proofs up to isomorphism.

The file `Algebra.Structures` in the standard library [DDG⁺19] contains definitions for monoids and other algebraic structures as record types. Because of the number of laws that need to be supplied to define a monoid, the example was reduced to a custom definition of magmas together with a proof that the carrier type is a set, see fig. 3.6.

A path between s_1 and s_2 coincides at time i with a “point-wise defined” algebraic structure, a magma. So defining such path is equivalent with defining a magma at every time i that coincides with the given s_1 and s_2 on its endpoints. To understand how

```

op2 : Op2 N0
op2 (x , p) (y , q) =
  (predN (x + y) , (predN (predN (x + y)) , sumLem x y p q) )

s2 : myMagma _ _
s2 = record {
  Carrier = N0 ;
  _◊_ = op2 ;
  s = transport (λ i → isSet (fEq i)) isSetN
}

```

Figure 3.6: The implementation of the definition of s_2 in agda works by giving a term for the record type magmas. Such a term, called a record, takes all the components of a magma such as the carrier set, operator and equivalence relation on the carrier set and groups them in one object.

```

zeroPath : (i : I) → (fEq i) ≡ (fEq i0)
zeroPath i = λ j → fEq (i ∧ (~ j))

...
op1' : PathP (λ i → Op2 (fEq i)) op1 op2'
op1' i x y = transport (sym (zeroPath i))
  (op1 (transport (zeroPath i) x) (transport (zeroPath i) y))

```

Figure 3.7: This definition of the intermediate operator is done by transporting the arguments along zeroPath back to fEq i0, applying op1. The result is then translated back forward to fEq i.

this works, look at the definition of magmas as record types. Record types are nested dependent sums and a path between two terms of a dependent sum is composed of paths between the components of the dependent sum. So for each component of the record type there should be path, including the carrier sets and operators on both structures.

It can be proven using the map $f : n \mapsto n + 1$ that there is an equivalence between \mathbb{N} and \mathbb{N}_0 which gives by the univalence axiom (proven to hold in def. 3.3.7 by section 3.4) a path $\mathbb{N} \equiv \mathbb{N}_0$. This gives a path between one component in the term of the record types of magmas.

Another component is the operator. In this case, the operator is actually a path between the two given operators. So we need a point-wise defined operator that coincides with the respective operators on its ending points. The construction is done by transporting the arguments of the operator from intermediate carrier set fEq i to the carrier set fEq i0 which is equal to \mathbb{N} and applying op1 there, see fig. 3.7.

The operator op1' from fig. 3.7 does not have the right operators at its ending points however and it is necessary to define another operator, see fig. 3.8.

The rest of the definition of the intermediate point-wise defined algebraic structure is straightforward. Once a term algPath of the equality $s_1 \equiv s_2$ (see fig. 3.9) between both magmas is obtained by assembling all point-wise constructs in a record type, algebraic properties of magmas can be transported.

For example, such a transport of properties could consist of mapping a term of the property that s_1 is commutative to a term of the property that s_2 is commutative, see

```

startLemma : op1 = op1'
startLemma i = λ x y → x + (transportRefl y) i

endLemma : op2' = op2
endLemma i x y =
  (op2' x y
   ≡⟨ refl ⟩
   transport fEq (op1 (transport (sym fEq) x)
                     (transport (sym fEq) y))
   ≡⟨ transpR
      (op1 (transport (sym fEq) x)
            (transport (sym fEq) y)) ⟩
   f (op1 (transport (sym fEq) x)
         (transport (sym fEq) y))
   ≡⟨ doubleCong (λ x y → f (op1 x y))
      (transpL x) (transpL y) ⟩
   f (op1 (g x) (g y))
   ≡⟨ cong f (l2' x y) ⟩
   f (predN (predN ((fst x) + (fst y))))
   ≡⟨ refl ⟩
   (suc (predN (predN ((fst x) + (fst y))))) ,
   (predN (predN ((fst x) + (fst y)))) , refl )
   ≡⟨ le x y ⟩
   (predN (fst x + fst y) ,
    (predN (predN (fst x + fst y)) ,
     sumLem (fst x) (fst y) (snd x) (snd y)))
   ≡⟨ refl ⟩
   op2 x y ■) i

pathLemma : (PathP (λ i → Op2 (fEq i)) op1' op2') =
  PathP ((λ i → Op2 (fEq i))) op1 op2
pathLemma = cong2 (PathP (λ i → Op2 (fEq i))) startLemma endLemma

opi : PathP (λ i → Op2 (fEq i)) op1 op2
opi = transport pathLemma op1'

```

Figure 3.8: This code shows two lemmas that prove that the endpoints of the original intermediate operator op_i' are equal to op_1 and op_2 in terms of the path identity type. Using these proofs and `transport`, a new intermediate operator op_i is defined that does satisfy the requirements for the intermediate operator of the intermediate magma.

```

isSeti : (i : I) → isSet (fEq i)
isSeti i = {!transport (λ j → isSet (fEq (i ∧ j))) isSetN!}

algPath : s1 = s2
algPath = λ i → record {
  Carrier = (fEq i) ;
  _◊_ = opi i ;
  s = isSeti i
}

```

Figure 3.9: A term of the intensional identity type between the two algebraic structures is a structure depending on a parameter i that has components depending on the parameter i . Each component was point-wise defined before and can be assembled into a point-wise structure.

```

isCommutative : myMagma → Set _
isCommutative m =
  (x y : (Carrier m)) → (x m._◊_ y) = y m._◊_ x

com1 : isCommutative s1
com1 = +-comm

com2 : isCommutative s2
com2 = transport (λ i → isCommutative (algPath i)) com1

```

Figure 3.10: A transport of proofs using the term of equality $s_1 = s_2$

fig. 3.10.

We could conclude at this point that if there is a homotopy type equivalences, (which is simply a bijection for discrete set-like types) between the carrier types of two terms of an algebraic type that have “compatible operations”, there is an equality between the terms. In this case, the equality is $s_1 = s_2$. The compatibility of the operations means that the algebraic operations can be connected by a path but actually corresponds to the homotopy type equivalence being a morphism in category theory.

An application of the path $s_1 = s_2$ between the two magmas, is that the proof of commutativity of the first magma s_1 can be transported to a proof of the commutativity of the second magma s_2 , defined as $com_2 = transport (λ i → isCommutative (algPath i)) com_1$. Although the property of being commutative is very simple, it can be generalized to more complicated properties and algebraic structures. This shows that it may be possible to do a kind of “isomorphism-invariant” algebra in practice. The complete code of this implementation can be found on [Van19] and verified with the installation instructions.

However, due to the numerous trivial lemmas that have to be implemented and the unusual homotopy theory inspired approach, it is not the easiest approach for proving algebraic properties. Definitional identities that in traditional mathematics simply hold by simple computation do not only hold up to some path in univalent type theory and cubical type theory. This means that simple facts about natural numbers are not simple anymore. However, not all structures in mathematics are as well-known

N.	Name	Day	Month	Year
4	John	30	5	1956
...
42	Sun	20	3	1980

N.	Name	Month	Day	Year
4	John	5	30	1956
...
42	Sun	3	20	1980

Table 3.1: The European and American databases on [Lic13b], p. 47.

as natural numbers and the structural and overly rigorous approach that such type theories require may be more beneficial in a general setting. At least in this case of simple traditional algebra, it seems easier to use specialized software such as [HBJ⁺18] to carry out algebraic computations and proofs.

3.5.2 Generic datatypes

Apart from mathematical applications, there are also more practical applications. Datatypes in mainstream programming languages such as Java can be parametrized by parameters. For example, a list may be parametrized by the type of its contents. The resulting datatype is then called generic, the approach is called *generic programming*. There might be multiple ways of implementing a datatype and several implementations might be equivalent. This is where the univalence axiom comes into play.

Assume we are given two date representations, “(month, day)” and “(day, month)”. These representations are equivalent by the swapping map. Univalence gives a path between the equivalent representations. Assume databases are generically defined for a certain date representation. For example, the type of databases is defined as $\text{List}(\mathbb{N} \times \text{String} \times (X \times \mathbb{N}))$ where $X = \mathbb{N} \times \mathbb{N}$ is any type of date representation. We would like to prove that the databases given in table 3.1 are equal.

The path between representations is lifted to a path on the level of the type of databases. The lifted path becomes a path between two different databases using two different but equivalent representations. Because paths behave like equality, the two databases are equal. The example has been implemented in the file `Cubical.Experiments.Generic` of [MV18].

This simple example shows that the implementation of univalence in cubical type theory allows a programmer to treat generic datatypes as the same datatype up to equivalence of the parameter type.

3.5.3 Formalizing algebraic topology

Higher inductive types

Higher inductive types are a concept coming from [VAC⁺13]. A *higher inductive type* is a type that has formation and construction rules as other types but these rules are accompanied by additional typing rules stating how to produce terms of the path types

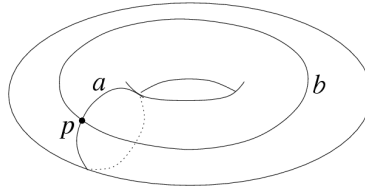


Figure 3.11: A topological representation of the torus [Din05].

between its terms. In other words, the type is generated by constructors and equalities. Higher inductive types are useful in synthetic homotopy type theory.

The first attempts at using higher inductive types in the cubical presheaf model date from [LB15], [Hub16a] and [CCHM16] but it was only until recently that the semantics were worked out in [LS17] and [CHM18].

Example 3.5.1. *The circle S^1 as a topological object is defined as line of which the endpoints are identified. In the language of higher inductive types this becomes two constructors: $\text{base} : S^1$, $\text{loop} : \text{base} = \text{base}$.*

Example 3.5.2. *The torus is defined similarly but now we have an extra loop which is introduced with an extra constructor. In addition, both loops intersect, meaning that doing one loop is topologically equivalent to talking the other loop. This is written as a path that starts and ends at the second loop: $\text{square} : \text{PathP } (\lambda i \rightarrow a \ i = a \ i) \ b \ b$. See fig. 3.11. Another way to construct the torus is given in [VAC⁺13], section 6.6.*

Algebraic topology

Homotopy groups are used in algebraic topology to compute the structure of topological spaces with the help of group theory. The first homotopy group of a topological space X is called the *fundamental group* of the space and denoted by $\pi_1(X)$ or $\Omega_1(X)$. For more information on general algebraic topology, see the text [Hat01].

The fundamental group of S^1 is formalized in `Cubical.HITs.S1.Base` as $\Omega(S^1) \equiv (\text{base} = \text{base})$ for some base points `base`. So the fundamental group is all the identifications of the base point with itself through a term of the path type. In the context of the univalence axiom, this path type can have multiple terms that represent paths. So the terms correspond to closed paths which are loops in S^1 as a topological space. The loops are generated by iterating the reflexive `loop` from the definition of S^1 as a higher inductive type. Because there are \mathbb{Z} ways of iterating this loop, it is no wonder that $\Omega_1(S^1) \simeq \mathbb{Z}$. General *homotopy groups* of types are defined by iterating the Ω construction. In formal homotopy type theory, general homotopy groups are denoted by $\Omega_i(X)$. These groups are quite complicated to compute for high values of i . The proofs use the Hopf fibration and long exact sequences. Many cases were already proven in homotopy type theory of in [VAC⁺13], Chapter 8. See also [Lic13a] for a survey of most results that have been proven in homotopy type theory and the methods used.

The following theorem was proven in formal homotopy type theory:

Theorem 3.5.3. $\Omega_4(S^3) \simeq \mathbb{Z}/\mathbb{Z}_2$

add information of im-
itations of higher induc-
types

The proof uses several advanced concepts such as for example the Blakers-Massey theorem, the James construction and Gysin exact sequences. See the full text of [Bru16].

A summary of the concepts used in this constructive proof compared to the proof in classical algebraic topology is given on [Bru16], page 123. The author mentions in the conclusion that cubical concepts are not always perfectly suited for this proof.

Nevertheless, there have been attempts at implementing the proof in cubical type theory based on the sketch in [Bru16], Appendix B. The first implementation was in the module `Examples.Brunerie` of the library [CCHM15] but is very long. The latest version is implemented in Agda and can be found in the file `Cubical.Experiments.Brunerie` of [MV18].

In these implementations, the definition of n in the expression $\Omega_4(S^3) \simeq \mathbb{Z}/n\mathbb{Z}$ is formalized as an Agda expression called `brunerie` or the “Brunerie number” that makes use of the univalence axiom. The expression is the application of a composition of several complicated functions to a representation of S^3 as a higher inductive type. Some of these functions use the concepts necessary for the proof of univalence but do not reference it explicitly. According to the proof of canonicity for natural numbers (see [Hub17] and [CHS19]), the expression `brunerie` should be equal to a numeral (that uses only a finite number of applications of `suc`), more specifically there should exist a term of the identity type

$$\text{Id}_{\mathbb{N}}(\text{brunerie}, \text{suc } (\text{suc } \text{zero}))$$

. Unfortunately the expression `brunerie` does not normalize to the number 2 in a finite amount of time. The length of shortest reduced form of `brunerie` was only reduced from a length in characters of order 10^6 to 10^5 according to [Bru18]. There have been many attempts to solve this problem during the last years but there remain two problems:

- Although the original expression of n in [CCHM15] has been rewritten in [MV18], it may contain errors. Besides, canonicity only implies that the result of reducing `brunerie` is possible in theory. The effective procedure for computing the number 2 is not given in [CHS19].
- The library Agda Cubical or Agda itself may contain bugs that prevent the proof from generating terms that are valid in the cubical set model.

Recently there has been a lot of ongoing activity in the development of the Agda Cubical library and `buffers` may have eliminated some reasons for the lack of normalization of the expression. Although this example still has problems, it shows how univalence could be used to formalize proofs of algebraic topology and verify their correctness, at least for relatively lower order homotopy groups.

3.5.4 Future applications

Until now, most applications of cubical type theory are related to implementing simple inductive types and higher inductive types. Most mathematical concepts from univalent type theory [VAC⁺13], part 2 are built with higher inductive types. So it should be possible to formalize more advanced concepts in cubical type theory. Concepts from univalent type theory that may be useful and have not been implemented with the help of cubical type theory and [MV18] include:

- The *structure identity principle* defined in [VAC⁺13], theorem 9.8.2 and [Acz12] states that isomorphic structures are equal. As a consequence proofs can be transferred or recycled between isomorphic structures. This principle could also be seen as a generalization of the simple example in section 3.5.1 to any category. The principle has not been formalized in all generality yet in cubical type theory but if this was done, it would be much easier for a practicing mathematician to identify isomorphic structures, one of the “main benefits” of homotopy type theory. Right now, software such as [HBJ⁺18] has to be used to reason about isomorphic algebraic structures.
- The application in section 3.5.3 needs the use of the univalence axiom. The theory of algebraic topology has some other theorems that can be formulated in homotopy type theory and have proofs using univalence such as Freudenthal, Blakers-Massey and Seifert-Van Kampen. These theorems can be used to compute homotopy groups in algebraic topology by putting information about smaller parts of a space together. The proofs have been formulated in the programming language Agda as a library [HBC⁺18] based on the theorems that were more informally presented in [VAC⁺13], chapter 8. For now, section 3.5.3 shows that computations do not succeed for the fourth fundamental group of S^3 , but it may be possible that computations of other homotopy groups of topological spaces are more efficient when [HBC⁺18] is combined with an interface to cubical type theory, for example through [MV18].
- Next to homotopy groups, there are also homology groups of topological spaces. For a fixed topological space, these do not necessarily have to be the same as the homotopy groups. Recently, synthetic homology groups have been studied in homotopy theory with filling operations of cubical type theory in [Gra18a], based on the initial approach to synthetic homotopy theory in [LB15]. The results in [Gra18a] have not been completely formalized in cubical type theory because the reasoning with higher-dimensional paths in cubical type theory posed some challenges. There exist computational systems for computing homology (and homotopy) groups that do not make use of type theory such as [GSS⁺99] and it might be interesting to check how [Gra18a] improves on this. Type theory may be better suited at solving complex new problems in computational topology than just trying to do this with the algorithms in [GSS⁺99].

A list of open problems for univalent type theory is on [SRvD⁺19] and contains subjects such as semantics, some of which have been solved with cubical type theory. There are also other open problems that remain relevant for future improvements on cubical type theory because cubical type theory is an extension of univalent type theory.

3.6 Alternative cubical models of univalence

There are several variations of *def.* 3.3.7.

3.6.1 Nominal models

The first models of cubical type theory in category theory such as for example the model defined in [BCH14] did not use the lattice structure as in def. 3.1.1 but required complicated Kan extension properties that also resulted in more complicated proofs of univalence such as the one in [BCH18]. They are called *monoidal* or sub-structural because their structure can be represented using only products of an interval object. The elements of the base category are independent objects representing dimensions, an approach used in [Hub16a] that is inspired by the theory of nominal sets which is introduced in [Pit13] as explained in [Pit14].

The lattice structure that is introduced in def. 3.1.1 was historically later introduced by [Hub16a] for simplifying composition and using higher inductive types. The unfinished text [Awo16] gives an overview of different variations of cubical categories that can be used to model cubical type theory.

3.6.2 Cartesian computational type theory

Computational type theory

The American approach is called “cartesian cubical computational type theory” and discussed in [ABC⁺17], [AHH18]. This version of cubical type theory is based on an extensional variant of intuitionistic type theory. This means that the identity type is not the intensional identity type of def. 1.2.5. This identity type, denoted by “ \equiv ”, is more like “equality by definition” and called the extensional identity type. It also satisfies the uniqueness of identity proof principle. Another identity type is introduced on top of the theory around the intensional identity type. This is also the approach taken by NuPRL and computational type theory on which cartesian cubical computational type theory is based. Using the extensional identity type, these theories extend type theory with quotients of types, set comprehension types, partial recursive function types, intersection types and other special types [Con11], p. 6. Computational type theory has been implemented in the proof assistant NuPRL and the underlying formal theory was described using partial equivalence relations (PERs) in [All87] according to [AHH18] which was later generalized in [Har91]. A historical overview of the relation between this kind of type theory and def. 1.2.9 is given in [Con03] and recently [Con15]. A computational type theory is usually built as follows:

- First, there is a definition of a set of “closed terms” which are like the terms of the intensional type theories but they are untyped, in the sense that the introduction of types comes later. If M and N are closed terms, then new closed terms are formed by constructors such as

$$\lambda x.M, M(N), \text{fst}(M), \text{snd}(M), 0, \mathbb{M}, \dots$$

- There are operational semantics defined on the closed terms to specify how they compute. Operational semantics describe how evaluation of terms influences evaluation of other terms, evaluation of the term M to M_0 is denoted by $M \mapsto M_0$. The expression $M \Downarrow M'$ denotes that both terms M and M' reduce to the same normal form. Examples of operational semantics are given on [Har91], p. 75.

- A type A is introduced as a partial equivalence (transitive and symmetric), called a *PER* relation denoted by $\llbracket A \rrbracket$ over the terms. Then the term a is of type A if and only if $\llbracket A \rrbracket(a, a)$, also denoted $a \in A$. In that case a is called a *canonical value* of type A . So types are defined by all the closed terms that are reflexive according to a particular PER.
- Every type A has a special kind of identity for terms a, b defined by setting $a \doteq b$ if and only if $\llbracket A \rrbracket(a, b)$. In this case, a, b are said to have *equal canonical values*.

Cubical extensions

Cartesian cubical computational type theory is based on results in the technical report [AHH17] and [AHH18]. The more informal text [Ben18] gives an introduction in the style of [VAC⁺13]. The base category of cartesian cubical computational type theory not explicitly mentioned because dimension variables are not considered as independent objects. This implies the theory is not proven to be consistent in a presheaf model which was done for the cubical type theory introduced in earlier chapters of this text in [Hub16a], part 2. Instead, dimension names such as x, y, z are treated as a kind of variables that can be substituted by *dimension terms* r , where r is either an endpoint of the unit interval (0 or 1) or a dimension name. This makes substitutions possible of the form (x/y) that can be considered as diagonals in a square. This can also be seen as adding terms of the form $\varphi \equiv (x = y)$ to the face type \mathbb{F} of cubical type theory [Hub16a]. This is where the name “cartesian” comes from. Dimension names are grouped in *dimension contexts* Ψ and types are PERs that are indexed by dimension contexts. Opposed to the theory in previous chapters (based on [Hub16a]), there is no lattice structure on the dimension variables occurring in a dimension context Ψ , which means that substitutions $A(i/i \wedge j)$ do not make sense anymore.

Operations from cubical type theory that are crucial for applications use the lattice structure such as the filling operation given in def. 3.3.5. According to [Mö18c], the solution is to give a different composition operation, a stronger one. This operation is defined in [AHH18], section 4 or [Mö18c], p.18. This new composition operation generalizes the filling and composition operation. It is however split up into two operations: homogeneous composition hcomp and coercion coe . According to [Mö18c], this splitting, makes certain implementations shorter.

Univalence

There is also an intensional path equality type defined in [AHH18] as is done in [Hub16a]. To turn equivalences between types of some universe into paths between those types and prove the univalence axiom, there are two possibilities in cartesian cubical computational type theory:

- Introduce Glue types as in section 3.4, after [Hub16a]. This was also taken in a previous version of cartesian cubical computational type theory [ABC⁺17].
- In [AHH18], the Glue type is weakened to a type called v-type with simple typing rules. On top of the v-type , instances of the hcomp and coe composition operations can be defined.

Together with the extensional equality that comes with all computational type theories and the proof of univalence, this turns cartesian cubical computational type theory into the first *two-level* homotopy type theory. Other such two-level type theories did not, but this theory satisfies *canonicity of booleans* according to [AHH18]. This canonicity is however equivalent to canonicity of numerals (see def. 3.3.8). The theory was also extended to include definitions for higher inductive types in [CH18] and [CH19]. The biggest advantage of this theory seems according to [Mö18c] to be that the proof of univalence with \mathbb{V} -types may be more efficient and allow an easier computation of the Brunerie number that was introduced in section 3.5.3.

Implementations

It was implemented in Haskell [MA18] and in the language RedPRL [SHA⁺18] with a proof of univalence that is slightly longer than the one in [MV18]. Development was continued in RedTT [SHC⁺18] which has many similarities but also extra features such as extension types according to [Mö18a] but stopped by the end of 2018. The library code in [SHC⁺18] also contains an alternative proof of univalence. The syntax is different from [MV18] but the structure of the libraries is very similar. The composition and completion operations used in [SHC⁺18] have also been implemented in [MV18] in the file `Foundations.CartesianKanOps`.

about the relational
ics, non-fibrant types?

3.6.3 Open challenges

The type theory in def. 1.2.9, which was described in [Ml75] has the following is *strongly normalizing*: there is no infinite sequence of reductions. This implies *decidability of type checking*, the procedure of checking whether a term does inhabit a certain type, but has not been proven at all for cubical type theory. It is noted in [CCHM16] that this could be done with resizing rules. The method described in [Abe13] may be useful.

The *confluency* of terms means that two sequences of reductions starting from the same term can be made to converge again. It is mentioned in the context of homotopy type theory in [HC11]. Confluency always holds in a strongly normalizing type theory but has also not been proven yet. An even weaker property is the property of *canonicity*. It says that every term can be written (uniquely) as a reduced form, containing only a finite amount of applications of the constructors of its type. Canonicity does not imply strongly normalizing and canonicity of numerals fails in homotopy type theory but is proven for (large parts of) cubical type theory in [Hub17, CHS19, Hub19]. This kind of canonicity is equivalent with the canonicity of booleans that was proven in [AHH18]. However, important benchmarks such as section 3.5.3 do not seem to benefit from the proof of canonicity due to the lack of an effective procedure for computing the reduced form. Canonicity for general types has not been proven for cubical type theory but is usually in other type theory not proven as well because it usually does not even hold for function types. There seem to be problems with the current implementations of \mathbb{G} types and there is ongoing work on trying to replace them with more efficient alternatives such as \mathbb{V} -types, see section 3.6.2 but cubes may not be a perfect solution according to [Mö18b].

Conclusion

This text gave an overview of the motivations and concepts behind the cubical set model of an extension univalent type theory in section 2.4 and the type theory called cubical type theory that is modelled by with cubical sets in chapter 3. Using the properties of the model, more precisely the `Glue` construction, it is possible to encode a proof of the univalence axiom in cubical type theory, see section 3.4. The application of the proof of univalence in cubical type theory to isomorphic algebraic structures in section 3.5.1 was not immediate but allowed to study transport of properties of structures. In standard univalent type theory it would not be possible to compute the transports of such properties but with the help of cubical type theory this became possible. Studying these formalized proofs in cubical type theory may give a (better) topological intuition about using the structure identity principle (see [Acz12]) in computer assisted proofs. Having this intuition becomes very useful with the implementation of cubical type theory because it allows to improve and understand the theorems from homotopy type theory using the primitives of the cubical set model which are mainly cubes. However, there remain big challenges such as proving normalizability and using canonicity for in formalizing fundamental theorems in homotopy theory, see section 3.6.3. There is ongoing research whether the primitives of cubical type theory can be replaced by more general or efficient alternatives.

Which parts of [Shi18] are interesting? This remark was also made by other learners of homotopy type theory as [Boo18]

Index

- ∞ -groupoid, 17
- axiom K, 28
- base category, 23
- canonical value, 57
- canonicity, 58
- canonicity of booleans, 58
- canonicity of numerals, 43
- categories with families, 20
- category of elements, 24
- category with families, 20
- circle, 53
- coherence laws, 17
- composition structure, 36
- concatenation, 11
- confluency, 58
- connections, 29
- constructor, 3
- context, 3
- context category, 20
- context extension, 4, 21
- context morphism, 21
- context restriction by extent φ , 33
- contractible, 14
- cubical set, 31
- cubical type theory, 42
- Curry-Howard correspondence, 5
- De Morgan algebra, 29
- decidability of type checking, 58
- dependent product, 8
- dependent sum, 8
- dependent type, 22
- dependent types, 8
- dimension context, 57
- dimension term, 57
- eliminator, 3
- empty context, 20
- equal canonical values, 57
- equivalence, 14, 41
- equivalence induction, 45
- equivalence of categories, 25
- extensible, 34
- extent, 31, 34
- face lattice, 31
- face maps, 31
- families of sets, 21
- fibrant type, 37
- filling, 40
- formal type theory, 4
- formation rule, 3
- free De Morgan algebra, 30
- fundamental group, 53
- generic programming, 52
- groupoid, 17
- h-level, 17
- h-prop, 12
- h-set, 12
- higher inductive type, 52
- homotopy equivalence, 46
- homotopy fiber, 14
- homotopy groups, 53
- homotopy hypothesis, 20
- homotopy of paths, 19
- homotopy type theory, 15, 18
- implicit arguments, 4
- indexed, 7
- indiscernability of identicals, 10
- inhabited, 2
- intensional identity type, 8, 10
- interval object, 32
- involution, 29
- isomorphism, 46
- judgment, 2

- Kan extension property, 36
- Kan type, 37
- language of toposes, 48
- membership judgment, 3
- monoidal, 56
- natural deduction, 4
- normalizability, 10
- operations, 36
- partial element of extent φ , 34
- path induction, 10
- path type, 38
- PER, 57
- presheaf, 22
- presheaf model, 23
- program extraction, 7
- quasi-inverse, 46
- reflective subcategory, 23
- reflexitivity, 10
- reflexive graph, 23
- restriction, 22
- set of names, 30
- sheaves, 23
- simple types, 5
- strongly normalizing, 58
- structure identity principle, 55
- substitutions, 21, 30
- sum type, 5
- system, 35
- term, 2
- terms, 21
- torus, 53
- transitivity, 39
- two-level, 58
- type, 2
- type theory, 10, 11
- types, 21
- typing rules, 3, 4
- uniqueness of identity proofs principle, 15
- univalence, 13, 44
- univalent, 15
- univalent type theory, 13, 15
- universal extension property, 21
- universe, 7

Bibliography

- [ABC⁺17] Carlo Angiuli, Guillaume Brunerie, Thierry Coquand, Kuen-bang Hou, Robert Harper, and Daniel R. Licata. Cartesian cubical type theory. 2017.
- [Abe13] Andreas Abel. Normalization by evaluation: Dependent types and impredicativity. March 2013. Available on <http://www2.tcs.ifi.lmu.de/~abel/habil.pdf>.
- [ACD⁺19] Andreas Abel, Jesper Cockx, Nils Anders Danielsson, Ulf Norell, Andrés Sicard-Ramírez, et al. Agda 2.5.4.2: dependently typed programming language / interactive theorem prover, April 2019. Available on <https://github.com/agda/agda>.
- [Acz12] Peter Aczel. Homotopy type theory and the structure identity principle. <https://www.newton.ac.uk/files/seminar/20120207160016301-153011.pdf>, February 2012.
- [ADH⁺19] Emilio Jesús Gallego Arias, Maxime Dénès, Hugo Herbelin, Pierre-Marie Pédro, Matthieu Sozeau, Enrico Tassi, et al. Coq 8.9.0: a formal proof management system, January 2019. Available on <https://github.com/coq/coq>.
- [AHH17] Carlo Angiuli, Kuen-Bang Hou, and Robert Harper. Computational higher type theory iii: Univalent universes and exact equality. December 2017.
- [AHH18] Carlo Angiuli, Robert Harper, and Kuen-Bang Hou. Cartesian cubical computational type theory: Constructive reasoning with paths and equalities. In *27th EACSL Annual Conference on Computer Science Logic (CSL 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. Available on <https://www.cs.cmu.edu/~rwh/papers/cartesian/paper.pdf>.
- [All87] Stuart Allen. A non-type-theoretic definition of martin-löf’s types. April 1987.
- [Alt17] Thorsten Altenkirch. *Introduction to Homotopy Type theory, lecture notes*. 2017.
- [Awo16] Steve Awodey. Notes on cubical models of type theory. June 2016. Available on <https://github.com/awodey/math/blob/master/Cubical/cubical.pdf>.
- [Bau13] Andrej Bauer. Socio-technological aspects of making the hott book, 2013. Available on <https://www.ias.edu/ideas/2013/bauer-hott-book>.

- [BCdMH18] Edwin Brady, David Thrane Christiansen, and Jan de Muijnck-Hughes. Idris 1.3.1: Code generation. Available on <http://docs.idris-lang.org/en/latest/tutorial/miscellany.html#c-target>, April 2018.
- [BCH14] Marc Bezem, Thierry Coquand, and Simon Huber. A model of type theory in cubical sets. In *19th International Conference on Types for Proofs and Programs (TYPES 2013)*, volume 26, page 107–128, 2014.
- [BCH18] Marc Bezem, Thierry Coquand, and Simon Huber. The univalence axiom in cubical sets. *Journal of Automated Reasoning*, June 2018.
- [BD77] Raymond Balbes and Philip Dwinger. Distributive lattices. 1977.
- [Ben18] Bruno Bentzen. Cubical informal type theory: the higher groupoid structure. *arXiv*, Jun 2018.
- [BG70] De Bruijn and Nicolaas Govert. The mathematical language automath, its usage, and some of its extensions. In *Symposium on automatic demonstration*, page 29–61. Springer, 1970.
- [Bib19] Curry–Howard correspondence - Wikipedia, April 2019. [Online; accessed 12. Apr. 2019].
- [Boo18] Auke Booij. What i wish i knew when learning hott, 2018. Available on <https://abooij.github.io/wiwikwlhott/index.html>.
- [Bro05] Luitzen Egbertus Jan Brouwer. Life, art and mysticism, 1905. Available on https://projecteuclid.org/download/pdf_1/euclid.ndjfl/1039886518.
- [Bru16] Guillaume Brunerie. *On the homotopy groups of spheres in homotopy type theory*. PhD thesis, 2016. Available at <https://arxiv.org/pdf/1606.05916.pdf>.
- [Bru18] Guillaume Brunerie. Experiments in cubical type theory. Slides from a talk given at UCSanDiego available at <https://guillaumebrunerie.github.io/pdf/cubicalexperiments.pdf>, January 2018.
- [BV06] John Michael Boardman and Rainer M. Vogt. *Homotopy invariant algebraic structures on topological spaces*, volume 347. Springer, 2006.
- [CCHM15] Thierry Coquand, Cohen Cyril, Simon Huber, and Anders Mörtberg. Cubicaltt: Experimental implementation of cubical type theory. GitHub repository available at <https://github.com/mortberg/cubicaltt>, February 2015.
- [CCHM16] Cyril Cohen, Thierry Coquand, Simon Huber, and Anders Mörtberg. Cubical type theory: a constructive interpretation of the univalence axiom. *arXiv preprint arXiv:1611.02108*, 2016.
- [CD13] Thierry Coquand and Nils Anders Danielsson. Isomorphism is equality. *Indagationes Mathematicae*, 24(4):1105–1120, 2013.

- [CFL05] Luíz Cruz-Filipe and Pierre Letouzey. A large-scale experiment in executing extracted programs. 2005.
- [CH18] Evan Cavallo and Robert Harper. Computational higher type theory iv: Inductive types. January 2018.
- [CH19] Evan Callo and Robert Harper. Higher inductive types in cubical computational type theory. <https://www.cs.cmu.edu/~rwh/papers/higher/paper.pdf>, January 2019. (Accessed on 05/12/2019).
- [CHM18] Thierry Coquand, Simon Huber, and Anders Mörtberg. On higher inductive types in cubical type theory. *arXiv preprint arXiv:1802.01170*, 2018.
- [CHS19] Thierry Coquand, Simon Huber, and Christian Sattler. Homotopy canonicity for cubical type theory. February 2019. Available on <https://arxiv.org/abs/1902.06572>.
- [CM19] Evan Cavallo and Anders Mörtberg. A unifying cartesian cubical type theory. 2019. Available on <http://www.cs.cmu.edu/~ecavallo/works/unifying-cartesian.pdf>.
- [Con03] Robert L. Constable. Naïve computational type theory. <http://nuprl.org/documents/Constable/naive.pdf>, Feb 2003. (Accessed on 04/28/2019).
- [Con11] Robert L. Constable. The triumph of types: Creating a logic of computational reality. 2011. Available on <https://ncatlab.org/nlab/files/ConstableTriumphOfTypes.pdf>.
- [Con15] Robert L. Constable. Two lectures on constructive type theory. <https://www.cs.uoregon.edu/research/summerschool/summer15/notes/OPLSS-Short-2015-2.pdf>, jul 2015. (Accessed on 04/28/2019).
- [Coq13] Thierry Coquand. Type theory and univalent foundation, October 2013. Available on <http://www.cse.chalmers.se/~coquand/russell.pdf>.
- [Coq15] Thierry Coquand. A category of cubical sets. September 2015. Available on <http://www.cse.chalmers.se/~coquand/vv.pdf>.
- [Cur34] Haskell Brooks Curry. Functionality in Combinatory Logic. *PNAS*, 20(11):584, November 1934.
- [Dan12] Nils Anders Danielsson. Isomorphism implies equality, September 2012. Available at <https://homotopytypetheory.org/2012/09/23/isomorphism-implies-equality/>.
- [DDG⁺19] Nils Anders Danielsson, Matthew Daggitt, Guillaume Gallais, Ulf Norell, Nicolas Pouillard, et al. The agda standard library. Available at <https://github.com/agda/agda-stdlib>, 2019.
- [Din05] Jonathan Dinkelbach. Fundamental group of torus, February 2005. Available on https://commons.wikimedia.org/wiki/File:Fundamental_group_torus2.png.

- [Eil39] Samuel Eilenberg. On the relation between the fundamental group on a space and the higher homotopy groups. *Fundamenta Mathematicae*, 32(1):167–175, 1939.
- [Esc19] Martín Hötzel Escardó. Introduction to univalent foundations of mathematics with agda. <https://www.cs.bham.ac.uk/~mhe/HoTT-UF-in-Agda-Lecture-Notes/index.html>, May 2019.
- [Eur13] Academia Europea. Member page: Per martin-löf. Available on http://www.ae-info.org/ae/User/Martin-L%C3%B6f_Per, 2013.
- [FT63] Walter Feit and John Thompson. Chapter i, from solvability of groups of odd order, pacific j. math, vol. 13, no. 3 (1963. *Pacific journal of mathematics*, 13(3):775–787, 1963.
- [GAA⁺13] Georges Gonthier, Andrea Asperti, Jeremy Avigad, Yves Bertot, Cyril Cohen, and Et Al. A machine-checked proof of the odd order theorem, 2013.
- [Gen35] Gerhard Gentzen. Untersuchungen über das logische schließen. i. *Mathematische Zeitschrift*, 39(1):176–210, December 1935.
- [Gra17] Daniel R. Grayson. Vladimir voevodsky, 4 june 1966 – 30 september 2017. Memorial webpage at <http://www.math.ias.edu/Voevodsky/>, 2017.
- [Gra18a] Robert Graham. Synthetic homology in homotopy type theory. <https://arxiv.org/pdf/1706.01540.pdf#page=20&zoom=100,0,701>, December 2018. (Accessed on 05/14/2019).
- [Gra18b] Daniel R. Grayson. An introduction to univalent foundations for mathematicians. *Bulletin of the American Mathematical Society*, 55(4):427–450, March 2018.
- [GSS⁺99] Julio Rubio Garcia, Francis Sergeraert, Yvon Siret, et al. Kenzo: a symbolic software for effective homology computation. April 1999.
- [Har91] Robert Harper. Constructing type systems over an operational semantics. <https://core.ac.uk/download/pdf/82210261.pdf>, November 1991. (Accessed on 05/14/2019).
- [Hat01] Allan Hatcher. *Algebraic Topology*. Cambridge University Press, 2001. Available on <http://www.math.cornell.edu/~hatcher/AT/ATpage.html>.
- [HBC⁺18] Kuen-Bang Hou, Guillaume Brunerie, Evan Cavallo, Andrej Bauer, Guillaume Brunerie, et al. Homotopy type theory in agda: Development of homotopy type theory in agda. October 2018.
- [HBJ⁺18] Max Horn, Reimer Behrends, Christopher Jefferson, Markus Pfeiffer, Alexander Kononov, Alexander Hulpke, Steve Linton, et al. Gap - groups, algorithms, programming - a system for computational discrete algebra. Downloads available at <https://www.gap-system.org/>, November 2018.

- [HC11] Simon Huber and Thierry Coquand. Towards a computational justification of the axiom of univalence. September 2011.
- [Hey30] Arend Heyting. Die formalen regeln der intuitionistischen logik, 1930.
- [Hof97] Martin Hofmann. Syntax and semantics of dependent types. In *Extensional Constructs in Intensional Type Theory*, page 13–54. Springer, 1997.
- [Hof14] Pieter Hofstra. Presheaves of sets. Online notes available at <http://mysite.science.uottawa.ca/phofstra/MAT5147/presheaves.pdf>, 2014.
- [HS98] Martin Hofmann and Thomas Streicher. The groupoid interpretation of type theory. *Twenty-five years of constructive type theory (Venice, 1995)*, 36:83–111, 1998.
- [Hub16a] Simon Huber. *Cubical Interpretations of Type Theory*. PhD thesis, University of Gothenburg, Gothenburg, Sweden, November 2016. Available on <http://www.cse.chalmers.se/~simonhu/misc/thesis.pdf>.
- [Hub16b] Simon Huber. Cubical interpretations of type theory: Phd defense. November 2016. Available on <http://www.cse.chalmers.se/~simonhu/slides/defense.pdf>.
- [Hub17] Simon Huber. Canonicity for cubical type theory. *Journal of Automated Reasoning*, page 1–38, 2017.
- [Hub19] Simon Huber. Homotopy canonicity for cubical type theory. February 2019.
- [Jac17] Konrad Jacobs. Nicolaas govert de bruijn in oberwolfach. Available at https://opc.mfo.de/detail?photo_id=541, 2017.
- [Kan55] Daniel M. Kan. Abstract homotopy. *Proceedings of the National Academy of Sciences*, 41(12):1092–1096, 1955.
- [KL12] Chris Kapulkin and Peter Lefanu Lumsdaine. The simplicial model of univalent foundations (after voevodsky). *arXiv preprint arXiv:1211.2851*, 2012.
- [LB15] Daniel R. Licata and Guillaume Brunerie. A cubical approach to synthetic homotopy theory. In *Proceedings of the 2015 30th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, page 92–103. IEEE Computer Society, 2015.
- [Lic13a] Daniel R. Licata. Homotopy theory in type theory: Progress report. May 2013. Available on <https://homotopytypetheory.org/2013/05/20/homotopy-theory-in-type-theory-progress-report/>.
- [Lic13b] Daniel R. Licata. Programming and proving in homotopy type theory, August 2013. Available on <http://dlicata.web.wesleyan.edu/pubs/l13jobtalk/l13jobtalk.pdf>.

- [LOPS18] Daniel R. Licata, Ian Orton, Andrew M. Pitts, and Bas Spitters. Internal universes in models of homotopy type theory. *arXiv preprint arXiv:1801.07664*, 2018.
- [LS17] Peter LeFanu Lumsdaine and Mike Shulman. Semantics of higher inductive types. <https://arxiv.org/abs/1705.07088>, may 2017. (Accessed on 05/18/2019).
- [Lum10] Peter Lefanu Lumsdaine. *Higher categories from type theories*. PhD thesis, PhD thesis, Carnegie Mellon University, 2010.
- [MA18] Anders Mörtberg and Carlo Angiuli. Yet another cartesian cubical type theory. <https://github.com/mortberg/yacctl/>, June 2018.
- [Ml75] Per Martin-löf. An intuitionistic theory of types: Predicative part. In H. E. Rose and J. Shepherdson, editors, *Logic colloquium '73: proceedings of the Logic Colloquium, Bristols*, volume 80 of *Studies in Logic and the Foundations of Mathematic*, page 73–118. Elsevier, Amsterdam, 1975.
- [MV18] Anders Mörtberg and Andrea Vezzosi. Cubical: An experimental library for cubical agda. Github repository, October 2018. Available on <https://github.com/agda/cubical>.
- [Mö18a] Anders Mörtberg. Cubical agda. Blog post available at <https://homotopytypetheory.org/2018/12/06/cubical-agda/>, December 2018.
- [Mö18b] Anders Mörtberg. Cubical variations: Hits, $\pi_4(s^3)$ and yacctl. <http://www.cs.cmu.edu/~amoertbe/slides/Mortberg-MURI18.pdf>, March 2018. (Accessed on 05/18/2019).
- [Mö18c] Anders Mörtberg. Yet another cartesian cubical type theory. <http://www.cs.cmu.edu/~amoertbe/slides/MortbergBonn.pdf>, June 2018.
- [Nuy18] Andreas Nuyts. Presheaf models of relational modalities in dependent type theory. May 2018.
- [Ort19] Richard Ian Orton. *Cubical Models of Homotopy Type Theory-An Internal Approach*. PhD thesis, University of Cambridge, 2019. Text available at <https://www.repository.cam.ac.uk/handle/1810/289441> and code on <https://doi.org/10.17863/CAM.35681>.
- [Pal14] Erik Palmgren. Lecture notes on type theory, January 2014. Available on <http://staff.math.su.se/palmgren/lecturenotesTT.pdf>.
- [PCZF⁺18] Clément Pit-Claudel, Théo Zimmermann, Jim Fehrle, Maxime Dénès, et al. Coq 8.9.0: Reference manual, December 2018. Available on <https://coq.inria.fr/refman/addendum/extraction.html#a-detailed-example-euclidean-division>.
- [Pea79] Giuseppe Peano. *Arithmetices principia, nova methodo exposita*. *English translation in 1899*, page 83–97, 1879.

- [Pit13] Andrew M. Pitts. *Nominal Sets: Names and Symmetry in Computer Science*. Cambridge University Press, New York, NY, USA, 2013.
- [Pit14] Andrew M. Pitts. Nominal presentation of cubical sets models of type theory. 2014. Available at <http://drops.dagstuhl.de/opus/volltexte/2015/5498/pdf/12.pdf>.
- [Pit18] Andrew Pitts. Axiomatizing cubical sets models of univalent foundations, September 2018. Available on <https://hott-uf.github.io/2018/slides/PittsHoTTUF2018.pdf>.
- [PW14] Álvaro Pelayo and Michael Warren. Homotopy type theory and voevodsky’s univalent foundations. *Bulletin of the American Mathematical Society*, 51(4):597–648, 2014.
- [Rod18] Sigrid Roden. Prof. martin hofmann, phd. Message on the website of Ludwig-Maximilians-Universität München <https://www.tcs.ifl.lmu.de/people/martin-hofmann>, 2018.
- [Rus03] Bertrand Russel. *The principles of mathematics*. Cambridge University Press, 1903. Available on <http://fair-use.org/bertrand-russell/the-principles-of-mathematics/>.
- [S⁺11] Urs Schreiber et al. Sheafification. January 2011. Available on <https://ncatlab.org/nlab/show/sheafification>.
- [SCS⁺18] Urs Schreiber, David Corfield, Mike Schulman, David Roberts, Praphulla Koushik, and Others. Homotopy hypothesis. <https://ncatlab.org/nlab/show/homotopy+hypothesis>, 2018.
- [SHA⁺18] Jonathan Sterling, Kuen-bang Hou, Carlo Angiuli, Evan Cavallo, James Wilcox, et al. The people’s refinement logic. <https://github.com/RedPRL/sml-redprl>, 2018.
- [SHC⁺18] Jonathan Sterling, Kuen-bang Hou, Evan Cavallo, Carlo Angiuli, et al. Redtt: a core language for cartesian cubical type theory with extension types, November 2018. File available at <https://bit.ly/2FxsTy1>.
- [Shi18] Andy Shiue. Cubical type theory for dummies, April 2018. Available on <https://gist.github.com/AndyShiue/cfc8c75f8b8655ca7ef2ffeb8cfb1faf/>.
- [SRvD⁺19] Bas Spitters, David Roberts, Floris van Doorn, Mike Shulman, Steve Awodey, Ulrik Buchholtz, Urs Schreiber, et al. Open problems in homotopy type theory. <https://ncatlab.org/homotopytypetheory/show/open+problems>, May 2019.
- [Str06] Thomas Streicher. Identity types vs. weak ω -groupoids. Talk given at the Work- shop “Identity Types – Topological and Categorical Structure” (Uppsala, November 13-14, 2006), November 2006.
- [Swa14] Andrew Swan. An algebraic weak factorisation system on 01-substitution sets: A constructive proof. *arXiv preprint arXiv:1409.1829*, 2014.

- [Swa18] Andrew Swan. Identity types in algebraic model structures and cubical sets. <https://arxiv.org/abs/1808.00915>, August 2018. (Accessed on 05/12/2019).
- [VAC⁺13] Vladimir Voevodsky, Steve Awodey, Thierry Coquand, et al. Homotopy type theory: Univalent foundations of mathematics. *Institute for Advanced Study (Princeton)*, 2013.
- [VAG⁺10] Vladimir Voevodsky, Benedikt Ahrens, Daniel Grayson, et al. Unimath-a library of formalized mathematics. Github repository available at <https://github.com/UniMath/UniMath>, 2010.
- [Van19] Willem Vanhulle. Transport of magmas. May 2019. GitHub repository on <https://github.com/wvhulle/transport-magmas>.
- [Voe09] Vladimir Voevodsky. Notes on type systems. Unpublished note available at https://github.com/vladimirias/old_notes_on_type_systems/raw/master/old_notes_on_type%20systems.pdf, September 2009.
- [Voe10] Vladimir Voevodsky. The equivalence axiom and univalent models of type theory. Notes from a talk at Carnegie Mellon University available at http://www.math.ias.edu/vladimir/files/CMU_talk.pdf, 2010.
- [Voe14] Vladimir Voevodsky. Univalent foundations—new type-theoretic foundations of mathematics. Slides from a talk at IHP, Paris, on April 22, 2014, available at https://math.ias.edu/vladimir/files/2014_04_22_slides.pdf, 2014.
- [Voe16] Vladimir Voevodsky. Unimath - a library of mathematics formalized in the univalent style. Talk at ICMS, Berlin available at https://www.math.ias.edu/vladimir/sites/math.ias.edu.vladimir/files/2016_07_14_Berlin_ICMS_short.pdf, July 2016.
- [Wei16] Jonathan Weinberger. The cubical model of type theory. Master’s thesis, November 2016. Available on <https://jonathanweinberger.files.wordpress.com/2016/10/msc-weinberger.pdf>.

DEPARTMENT OF MATHEMATICS

Celestijnenlaan 200B
3001 LEUVEN, BELGIË
tel. + 32 16 32 70 15
<https://wis.kuleuven.be/english>

