A Higher Structure Identity Principle

Benedikt Ahrens
School of Computer Science
University of Birmingham
United Kingdom
b.ahrens@cs.bham.ac.uk

Paige Randall North
Department of Mathematics
The Ohio State University
United States
north.138@osu.edu

Michael Shulman
Department of Mathematics
University of San Diego
United States
shulman@sandiego.edu

Dimitris Tsementzis
Princeton University
Rutgers University
United States
dimitrios.tsementzis@gmail.com

Abstract

The ordinary Structure Identity Principle states that any property of set-level structures (e.g., posets, groups, rings, fields) definable in Univalent Foundations is invariant under isomorphism: more specifically, identifications of structures coincide with isomorphisms. We prove a version of this principle for a wide range of higher-categorical structures, adapting FOLDS-signatures to specify a general class of structures, and using two-level type theory to treat all categorical dimensions uniformly. As in the previously known case of 1-categories (which is an instance of our theory), the structures themselves must satisfy a local univalence principle, stating that identifications coincide with "isomorphisms" between elements of the structure. Our main technical achievement is a definition of such isomorphisms, which we call "indiscernibilities," using only the dependency structure rather than any notion of composition.

CCS Concepts: • **Theory of computation → Type theory**; *Logic and verification*; Constructive mathematics.

Keywords: homotopy type theory, univalent foundations, structure identity principle, categories, equivalence principle

1 Introduction

1.1 The Structure Identity Principle (SIP)

A fundamental logical principle is the "indiscernibility of identicals", i.e., equal objects have the same properties:

$$x = y \rightarrow \forall \text{ properties } P, (P(x) \leftrightarrow P(y)). \qquad (1)$$

However, properties invariant under weaker notions of sameness are also important. For instance, group-theoretic properties satisfy a similar principle for *isomorphic* groups:

$$G \cong H \rightarrow \forall \text{ group-theoretic properties } P, (P(G) \leftrightarrow P(H));$$

while category-theoretic properties are invariant even under *equivalence* of categories:

$$A \simeq B \rightarrow \forall \text{ category-theoretic properties } P, (P(A) \leftrightarrow P(B)).$$

The idea is summarized in Aczel's Structure Identity Principle (SIP) [1]: *Isomorphic (or equivalent) mathematical structures are structurally identical; i.e., have the same structural properties.* But it remains to characterize the "structural properties" for a given notion of structure. For instance, Blanc [5] and Freyd [7] devised a syntax for category-theoretic properties and showed that they are invariant under equivalence. And Makkai [9] introduced general notions of signature and equivalence for higher-categorical structures, along with a language for their properties called First Order Logic with Dependent Sorts (FOLDS), and proved that FOLDS-properties are invariant under FOLDS-equivalence.

1.2 The SIP in Univalent Foundations

Inspired by Makkai (see [12, p. 1279]), Voevodsky conceived Univalent Foundations (UF) with a similar but more ambitious goal: a foundational language for mathematics, *all* of whose constructions are invariant under equivalences of structures. Since proofs are particular constructions, this implies a similar invariance of properties.

The formal language of UF and the closely related Homotopy Type Theory (HoTT) is Martin-Löf type theory, with

types regarded as (higher) groupoids; see [11] for background and notation. Voevodsky's univalence principle

$$\text{univalence} : \Pi_{(x,y:\mathcal{U})}(x =_\mathcal{U} y \xrightarrow{\sim} x \simeq y) \qquad (2)$$

ensures that all properties of *types* are invariant under equivalences, since equivalences can be made into identifications (i.e., elements of the Martin-Löf identity type $x = y$):

$$\Pi_{(x,y:\mathcal{U})} \left(x \simeq y \rightarrow \Pi_{(P:\mathcal{U}\rightarrow\mathcal{U})} \left(P(x) \simeq P(y) \right) \right).$$

It was proven in [11, Section 9.9] and [6][1] that the same approach works for a wide range of mathematical structures. Both use a notion of "signature" to define general classes of structures and isomorphisms, and show (using (2) for the underlying types) that isomorphisms of structures are equivalent to identities of structures. Thus, indiscernibility of identicals implies indiscernibility of isomorphs.

1.3 SIPs for Categories and Higher Categories

The main restriction of the SIP of Section 1.2 is that it applies only to structures that naturally form a 1-category (or 1-groupoid). In particular, it excludes categories themselves. An SIP for categories is proved in [2, Theorem 6.17]: equivalence of (certain) categories is equivalent to identity of categories, yielding an analogous transport principle for categorical equivalences $x \simeq y$:

$$\Pi_{(x,y:\text{Cat})} \left(x \simeq y \rightarrow \Pi_{(P:\text{Cat}\rightarrow\mathcal{U})} \left(P(x) \simeq P(y) \right) \right).$$

However, to make this true, the categories themselves must satisfy a local univalence principle saying that isomorphism is equivalent to identity for their objects.

In the present paper, grown out of [10] and inspired by FOLDS, we generalize this to other (higher-)categorical structures. We give general notions of signature, structure, univalence, and equivalence, such that equivalence of univalent structures is equivalent to identity. Therefore, indiscernibility of identicals implies indiscernibility of equivalents:

$$\Pi_{(x,y:\text{uStr}(\mathcal{L}))} \left(x \simeq_\mathcal{L} y \rightarrow \Pi_{(P:\text{uStr}(\mathcal{L})\rightarrow\mathcal{U})} \left(P(x) \simeq P(y) \right) \right).$$

The primary difficulty is to give a suitable notion of isomorphism—which we also call *indiscernibility*—between *elements* of a general structure, so that we can define a structure to be univalent if this notion of isomorphism of its elements is equivalent to their identity. Our definition is a relativized form of the converse of (1), the *identity of indiscernibles*:

$$(\forall \text{ properties } P, (P(x) \leftrightarrow P(y))) \rightarrow (x = y). \qquad (3)$$

As long as we generalize "property" to "construction", the global form of this principle in HoTT/UF is trivially true because we have the haecceity $P(u) :\equiv (x = u)$. We get our notion of indiscernibility by relativizing this to a structure, allowing only constructions involving the data of a structure to which x, y belong. Thus, our univalent structures satisfy a local form of (3).

---

[1]The formalization of [6] compares the two independent results.

Remark 1.1. The SIP of Section 1.2 is equivalent to the statement that the 1-category *of* structures is univalent in the sense of [2]. This suggests a "Baez–Dolan microcosm principle" at work: the SIP *for* a given kind of structure should state that the (higher) category *of* such structures satisfies the necessary univalence principle for *its* SIP. We have not attempted to state or prove this precisely.

2 A Fresh Look at Univalent Categories

First we review [2] and [11, Chapter 9] with an eye to generalization. Both start by defining a **precategory** C as follows.

- A type C_0 of objects.
- For each $a, b : C_0$, a set $C(a, b)$ of morphisms.
- For each $a : C_0$, a morphism $1_a : C(a, a)$.
- For each $a, b, c : C_0$, a function

$$C(b, c) \rightarrow C(a, b) \rightarrow C(a, c).$$

- For each $a, b : C_0$ and $f : C(a, b)$, we have $f = 1_b \circ f$ and $f = f \circ 1_a$.
- For each $a, b, c, d : C_0$ and $f : C(a, b)$, $g : C(b, c)$, $h : C(c, d)$, we have $h \circ (g \circ f) = (h \circ g) \circ f$.

Note C_0 may not be a set, and for "large" precategories it almost never is. For instance, Set_0 is the type of sets, which by univalence is a proper 1-type. However, allowing arbitrary types of objects is problematic too. For instance, while the statement "a fully faithful and essentially surjective functor is an equivalence" in ZF is equivalent to the axiom of choice, for precategories in HoTT/UF it is generally false, even with the axiom of choice.

The solution is to impose a "local univalence" condition: for any $a, b : C_0$ there is a map $\text{idtoiso}_{a,b} : (a =_{C_0} b) \rightarrow (a \cong b)$, defined by path induction, and a precategory C is a **univalent category** if $\text{idtoiso}_{a,b}$ is an equivalence for all $a, b : C_0$. This implements the idea that "isomorphic objects are equal". Note that it implies that C_0 is a 1-type, since its identity types are all sets (0-types).

One then proves, using the univalence axiom, that this "local" form of univalence for objects of a category implies a "global" form of univalence for categories themselves:

Theorem 2.1 ([2, Theorem 6.17]). *For univalent categories C and $\mathcal{D}$, let $C \simeq \mathcal{D}$ be the type of categorical equivalences between C and $\mathcal{D}$; then*

$$(C = \mathcal{D}) = (C \simeq \mathcal{D}).$$

We will generalize this to other categorical structures, starting with a general vocabulary for expressing such things.

2.1 FOLDS-Signature for Categories

In [9], Makkai presents a definition of category in a language called First-Order Logic with Dependent Sorts (FOLDS). In contrast to HoTT/UF, FOLDS is not a foundational system for mathematics, but a kind of first-order logic designed for higher categorical structures. We will not use the logical

$$ci = di, ct_0 = dt_1, dt_0 = dt_2, ct_1 = ct_2, de_1 = de_2, ce_1 = ce_2$$
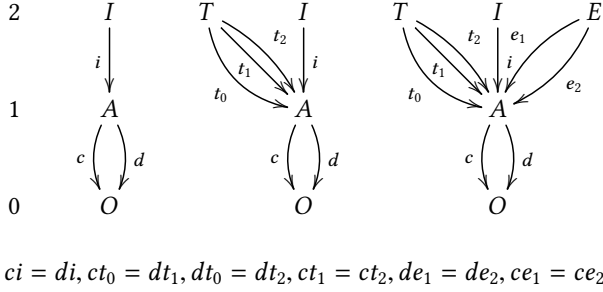
Figure 1. The FOLDS-signatures $\mathcal{L}_{\text{rg}}$, $\mathcal{L}_{\text{cat}}$, and $\mathcal{L}_{\text{cat+E}}$ (from left to right) for reflexive graphs, for categories, and for categories with equality predicate on arrows. The morphisms are subject to the indicated equalities.

syntax of FOLDS, but we adopt and generalize its notions of signature and structure.

A FOLDS-signature is an *inverse category with finite fan-outs*, whose objects are called *sorts*. The FOLDS-signature $\mathcal{L}_{\text{cat}}$ of categories is shown in Figure 1, along with the related FOLDS-signatures $\mathcal{L}_{\text{rg}}$ of reflexive graphs and $\mathcal{L}_{\text{cat+E}}$ of categories with equality (see Section 2.2). There are some relations on the composite arrows (e.g., the two composites $I \to A \rightrightarrows O$ are equal). The intent is that O is the sort of objects, A the sort of arrows, I the sort of identity arrows, and T the sort of composable pairs of arrows (with their composite; T stands for "triangle").

Using a set-theoretic metatheory, Makkai defined *structures* for a FOLDS-signature as certain functors into **Set**. In a dependently typed theory like HoTT/UF, however, it is more natural to interpret each sort as a dependent type indexed by the interpretations of all the sorts below it. For instance, in HoTT/UF a structure M for the FOLDS-signature $\mathcal{L}_{\text{cat}}$ in Figure 1 consists of

$$MO : \mathcal{U}$$
$$MA : MO \times MO \to \mathcal{U}$$
$$MI : \prod_{(x:MO)} MA(x,x) \to \mathcal{U}$$
$$MT : \prod_{(x,y,z:MO)} MA(x,y) \to MA(y,z) \to MA(x,z) \to \mathcal{U}$$

This forms the underlying data of a category: a type of objects, types of morphisms, and properties of "being an identity" and "being the composite". Later we will see that our univalence condition implies that the families MI and MT consist of propositions and MA consists of sets.

2.2 Axioms and Theories

To express which such structures are actually categories, Makkai introduced a logic over FOLDS-signatures, taking top-level sorts as "relation symbols". For instance, the axiom that any two composable arrows have a composite would be

$$\forall(x,y,z:O).\forall(f:A(x,y)).\forall(g:A(y,z)).$$

$$\exists(h:A(x,z)).T_{x,y,z}(f,g,h).$$

Any such axiom can be interpretated as a predicate on $\mathcal{L}$-structures in HoTT/UF. (Note that the interpretation of $\exists$ and $\vee$ involves propositional truncation.)

The axioms of a category also involve equality of arrows (though not of objects), e.g., the uniqueness of composites

$$\forall(x,y,z:O).\forall(f:A(x,y)).\forall(g:A(y,z)).\forall(h,h':A(x,z)).$$

$$T_{x,y,z}(f,g,h) \wedge T_{x,y,z}(f,g,h') \to (h = h'). \quad (4)$$

Just as in ordinary first-order logic, one can consider FOLDS either *with equality* or *without equality*. In the former, equality is only allowed between elements of sorts that are "one level below the top" like A. As usual, FOLDS with equality can be embedded in FOLDS without equality by adding equality relations to the signature, as with $\mathcal{L}_{\text{cat+E}}$ in Figure 1, along with axioms making them congruences (where all free variables should be considered to be universally quantified):

$$E_{x,y}(f,f)$$
$$E_{x,y}(f,g) \to E_{x,y}(g,f)$$
$$E_{x,y}(f,g) \wedge E_{x,y}(g,h) \to E_{x,y}(f,h)$$
$$E_{x,x}(f,g) \wedge I_x(f) \to I_x(g)$$
$$E_{x,y}(f,f') \wedge E_{y,z}(g,g') \wedge E_{x,z}(h,h') \wedge T_{x,y,z}(f,g,h)$$
$$\to T_{x,y,z}(f',g',h'). \quad (5)$$

A model M is **standard** if $ME_{x,y}$ is equivalent to the actual equality of $MA(x,y)$.[2] Like truncatedness, this will turn out to be a special case of our univalence condition.

2.3 FOLDS-Categories in Univalent Foundations

As noted above, in HoTT/UF we must consider what truncation level MO and $MA(x,y)$ should have. In a precategory, we require the types of arrows to be sets, suggesting the following analogous definition.

Definition 2.2. A **1-univalent FOLDS-category** M is

- A type $MO : \mathcal{U}$;
- A family $MA : MO \times MO \to \mathcal{U}$;
- A family $MI : \prod_{(x:MO)} MA(x,x) \to \mathcal{U}$;
- A family $MT : \prod_{(x,y,z:MO)} MA(x,y) \to MA(y,z) \to MA(x,z) \to \mathcal{U}$; and
- A family $ME : \prod_{(x,y:MO)} MA(x,y) \to MA(x,y) \to \mathcal{U}$,

such that

- Each type $MI_x(f)$, $MT_{x,y,z}(f,g,h)$, and $ME_{x,y}(f,g)$ is a proposition;
- Each type $MA(x,y)$ is a set;
- $ME_{x,y}(f,g) \leftrightarrow (f = g)$;

and the axioms of a category are satisfied.

Lemma 2.3. *The type of 1-univalent FOLDS-categories is equivalent to the type of precategories.*

---

[2]With reference to "identity of indiscernibles" from Section 1.3, standard equality amounts to adding haecceities into the structure explicitly.

Proof. The underlying data of MO and MA are the same. In one direction, let $MI_x(f) :\equiv (f = 1_x)$ and $MT_{x,y,z}(f, g, h) :\equiv (h = g \circ f)$. In the other, let 1_x be the unique $f : MA(x, x)$ with $MI_x(f)$, and $g \circ f$ the unique h with $MT_{x,y,z}(f, g, h)$. □

Convention 2.4. Below, we sometimes abuse notation by writing $x : O$ instead of $x : MO$, and similarly for the other sorts, when the particular structure M is clear from context.

Let us now consider how to define "univalent categories" using only the FOLDS-structure. The central problem is to characterize the type $(a \cong b)$ of isomorphisms in such a way as can be readily generalized to other signatures.

To start with, recall that by the Yoneda lemma, an isomorphism $\phi : a \cong b$ in a category C is equivalently a natural family of isomorphisms of sets $\phi_{x\bullet} : C(x, a) \cong C(x, b)$, where naturality in x means that $\phi_{y\bullet}(g) \circ f = \phi_{x\bullet}(g \circ f)$. In the language of FOLDS-categories the operation $\circ$ is replaced by the relation T, with a new variable h for the composite $g \circ f$:

- For each $x : O$, an isomorphism $\phi_{x\bullet} : A(x, a) \cong A(x, b)$; and
- For each $x, y : O$, $f : A(x, y)$, $g : A(y, a)$, and $h : A(x, a)$, we have $T_{x,y,a}(f, g, h) \leftrightarrow T_{x,y,b}(f, \phi_{y\bullet}(g), \phi_{x\bullet}(h))$.

This looks more promising, but it still privileges one of the variables of A over the other, and the relation T over I (and E). More natural from the FOLDS point of view is to give equivalences between hom-sets with a and b substituted into *all* possible "collections of holes":

For any $x : O$, an isomorphism $\phi_{x\bullet} : A(x, a) \cong A(x, b)$; (6)

For any $z : O$, an isomorphism $\phi_{\bullet z} : A(a, z) \cong A(b, z)$; (7)

An isomorphism $\phi_{\bullet\bullet} : A(a, a) \cong A(b, b)$. (8)

and similar logical equivalences between all possible "relations with holes":

$$T_{x,y,a}(f, g, h) \leftrightarrow T_{x,y,b}(f, \phi_{y\bullet}(g), \phi_{x\bullet}(h)) \tag{9}$$

$$T_{x,a,z}(f, g, h) \leftrightarrow T_{x,b,z}(\phi_{x\bullet}(f), \phi_{\bullet z}(g), h) \tag{10}$$

$$T_{a,z,w}(f, g, h) \leftrightarrow T_{b,z,w}(\phi_{\bullet z}(f), g, \phi_{\bullet w}(h)) \tag{11}$$

$$T_{x,a,a}(f, g, h) \leftrightarrow T_{x,b,b}(\phi_{x\bullet}(f), \phi_{\bullet\bullet}(g), \phi_{x\bullet}(h)) \tag{12}$$

$$T_{a,x,a}(f, g, h) \leftrightarrow T_{b,x,b}(\phi_{\bullet x}(f), \phi_{x\bullet}(g), \phi_{\bullet\bullet}(h)) \tag{13}$$

$$T_{a,a,x}(f, g, h) \leftrightarrow T_{b,b,x}(\phi_{\bullet\bullet}(f), \phi_{\bullet x}(g), \phi_{\bullet x}(h)) \tag{14}$$

$$T_{a,a,a}(f, g, h) \leftrightarrow T_{b,b,b}(\phi_{\bullet\bullet}(f), \phi_{\bullet\bullet}(g), \phi_{\bullet\bullet}(h)) \tag{15}$$

$$I_{a,a}(f) \leftrightarrow I_{b,b}(\phi_{\bullet\bullet}(f)) \tag{16}$$

$$E_{x,a}(f, g) \leftrightarrow E_{x,b}(\phi_{x\bullet}(f), \phi_{x\bullet}(g)) \tag{17}$$

$$E_{a,x}(f, g) \leftrightarrow E_{b,x}(\phi_{\bullet x}(f), \phi_{\bullet x}(g)) \tag{18}$$

$$E_{a,a}(f, g) \leftrightarrow E_{b,b}(\phi_{\bullet\bullet}(f), \phi_{\bullet\bullet}(g)) \tag{19}$$

for all $x, y, z, w : O$ and f, g, h of appropriate types. Fortunately, the additional data here are redundant. Since $\phi_{x\bullet}$, $\phi_{\bullet z}$, and $\phi_{\bullet\bullet}$ preserve identities and E is equivalent to identity by hypothesis, we obtain (17) to (19). Just as (9) means the $\phi_{x\bullet}$ form a natural isomorphism, (11) means the $\phi_{\bullet z}$ form a natural isomorphism, and (10) means these natural isomorphisms

arise from the same $\phi : a \cong b$. Given this, any one of Eqs. (12) to (14) ensures that $\phi_{\bullet\bullet}$ is conjugation by ϕ, and then the other two follow automatically, as do Eqs. (15) and (16). This suggests the following definition.

Definition 2.5. For a, b objects of a 1-univalent FOLDS-category, an **indiscernibility** from a to b consists of data as in Eqs. (6) to (8) satisfying Eqs. (9) to (16). We write $a \asymp b$ for the type of such indiscernibilities.

Theorem 2.6. *In any* 1-*univalent FOLDS-category, the type of indiscernibilities from a to b is equivalent to the type of isomorphisms $a \cong b$.* □

Definition 2.7. A 0-**univalent FOLDS-category** is a 1-univalent FOLDS-category such that for all $a, b : MO$, the canonical map $(a = b) \to (a \asymp b)$ is an equivalence.

Theorem 2.8. *A 1-univalent FOLDS-category is* 0-*univalent iff its corresponding precategory is a univalent category.* □

The point is that the definition of indiscernibility can be derived algorithmically from the FOLDS-signature for categories, by an algorithm which applies equally well to any FOLDS-signature. We will give this mechanism explicitly in Section 6. Then, for any $a, b : MK$ in some structure M, there will be a canonical map $(a =_K b) \to (a \asymp b)$, and we call M **univalent** if these are equivalences.

However, there are two mismatches between this example so far and the general theory we have just proposed. Firstly, we have assumed *ad hoc* that MA consists of sets. Secondly, we have just proposed that *all* sorts should satisfy a univalence property, but in the example of categories we have only considered this for the sort O. Fortunately, these two problems solve each other, and moreover remove the need to postulate "standardness" of equality.

Definition 2.9. A 2-**univalent FOLDS-category** M consists of the same type families MO, MA, MI, MT, ME as a 1-univalent FOLDS-category, such that ME is a congruence, each type $MI_x(f)$, $MT_{x,y,z}(f, g, h)$, and $ME_{x,y}(f, g)$ is a proposition, and the axioms of a category are satisfied with ME used in place of equality.

An indiscernibility between $f, g : A(a, b)$ in a 2-univalent FOLDS-category should consist of logical equivalences between instances of T, I, and E with f replaced by g in "all possible ways", clearly beginning with

$$T_{x,a,b}(u, f, v) \leftrightarrow T_{x,a,b}(u, g, v) \tag{20}$$

$$T_{a,x,b}(u, v, f) \leftrightarrow T_{a,x,b}(u, v, g) \tag{21}$$

$$T_{a,b,x}(f, u, v) \leftrightarrow T_{x,a,b}(g, u, v) \tag{22}$$

for all $x : O$ and u, v of appropriate types. But how do we put f in two or three of the places in T in the most general way? In Section 6 we will see that the answer is to assume an equality between objects and transport f along it.

Definition 2.10. For $f, g : A(x, y)$ in a 2-univalent FOLDS-category, an **indiscernibility** from f to g consists of the logical equivalences shown in Eqs. (20) to (30), for all $p : a = a$, $q : b = a$, and $r : b = b$.

$$T_{a,a,b}(q_*(f), f, u) \leftrightarrow T_{a,a,b}(q_*(g), g, u) \tag{23}$$

$$T_{a,b,b}(p_*(f), u, f) \leftrightarrow T_{a,b,b}(p_*(g), u, g) \tag{24}$$

$$T_{a,a,b}(u, r_*(f), f) \leftrightarrow T_{a,a,b}(u, r_*(g), g) \tag{25}$$

$$T_{a,a,b}((p,q)_*(f), r_*(f), f) \leftrightarrow T_{a,a,b}((p,q)_*(g), r_*(g), g) \tag{26}$$

$$I_a(q_*(f)) \leftrightarrow I_a(q_*(g)) \tag{27}$$

$$E_{a,b}(f, u) \leftrightarrow E_{a,b}(g, u) \tag{28}$$

$$E_{a,b}(u, f) \leftrightarrow E_{a,b}(u, g) \tag{29}$$

$$E_{a,b}((p,r)_*(f), f) \leftrightarrow E_{a,b}((p,r)_*(g), g) \tag{30}$$

Since T, I, and E are propositions, so is the type $f \asymp g$ of indiscernibilities. And $f \asymp f$, so by path induction we have $(f = g) \to (f \asymp g)$.

Theorem 2.11. *A 2-univalent FOLDS-category is 1-univalent iff the map $(f = g) \to (f \asymp g)$ is an equivalence for all f, g.*

Proof. Since $f \asymp g$ is a proposition, the latter condition implies that each $A(a, b)$ is a set. Thus, for "if" it suffices to show $E_{a,b}(f, g) \Rightarrow (f \asymp g)$, which holds since E is a congruence for T and I. For "only if", we must show $(f \asymp g) \Rightarrow (f = g)$ in a FOLDS-category. But since $E_{a,b}(f, f)$ always, $f \asymp g$ implies $E_{a,b}(f, g)$, hence $f = g$ by standardness. $\square$

Thus, by extending the "univalence" condition of a category from the sort O to the sort A, we encompass automatically the assumption that the hom-types in a precategory are sets and that the equality is standard.

Finally, we can even stop treating the top sorts specially.

Definition 2.12. A **FOLDS-category** consists of the same data and axioms as a 2-univalent FOLDS-category, but without the assumption that the types T, I, and E are propositions.

Now, the type $t \asymp t'$ of indiscernibilities between $t, t' : T_{x,y,z}(f, g, h)$ should consist of consistent equivalences between all types dependent on t and t'. But there are no such types in the signature, so $t \asymp t'$ is contractible. The same reasoning applies to I and E. Thus, the univalence condition for these sorts will assert simply that all of their path-types are contractible, i.e., that they are propositions.

Theorem 2.13. *A FOLDS-category is 2-univalent if and only if the canonical maps $(t = t') \to (t \asymp t')$, $(i = i') \to (i \asymp i')$, and $(e = e') \to (e \asymp e')$ are equivalences for all inhabitants of the types T, I, and E respectively.* $\square$

Thus, the notion of univalent category is determined only by the signature $\mathcal{L}_{\text{cat+E}}$, plus axioms (which are irrelevant for indiscernibilities and univalence). Our goal is to define notions of indiscernibility and univalence for any signature $\mathcal{L}$,

generalizing the theory of univalent categories to arbitrary higher-categorical structures.

3 Background: Two-Level Type Theory

In the following sections, we work in a two-level type theory (2LTT) as in [4], with axioms (M2) (Russell-style universes), (T1), (T2), and (T3) from [4, Section 2.4]. Building on [8], 2LTT is shown in [4, §2.5] to be modeled by simplicial sets.

2LTT has an "outer" (a.k.a. "strict") level, a Martin-Löf type theory with intensional identity types and uniqueness of identity proofs (UIP), and an "inner" level, a homotopy type theory with univalent universes. Both have their own—*prima facie* distinct—type formers $\Pi, \Sigma, +, \mathbf{1}, \mathbf{0}, \mathbb{N}$, intensional "=" with function extensionality, and universes. By axioms (T1) and (T2), we can **identify inner types with particular outer ones** so that Π, Σ, and $\mathbf{1}$ are "shared" between the levels [4, Lemma 2.11], so we need not distinguish those notationally. For other type constructors we annotate the outer variants with ^{s} (for "strict"), e.g., in $\mathbb{N}^s$. (In [4], the *inner* type formers are annotated.) We use the conventional typical ambiguity [11, Section 1.3] and hence refer to any universe by $\mathcal{U}$ (inner) resp. $\mathcal{U}^s$ (outer). We use $\equiv$ and $:\equiv$ to denote judgmental equality, e.g., in definitions.

We call a type A **fibrant** when it is isomorphic to an inner type A' (in the strict sense, modulo $\overset{\text{s}}{=}$). Fibrancy is structure rather than property, but following [4] we abuse language by talking about a type "being fibrant" for simplicity. Axiom (T3) states that **every fibrant type is inner**. Thus, fibrant types are closed under Π and Σ ([4, Lemma 3.5]).

A fibrant A has two identity types: for $a, b : A$ we have the strict identity type $a \overset{\text{s}}{=} b$ that satisfies UIP, and the homotopical identity type $a = b$ that is at the center of HoTT. We refer to elements of $a = b$ as "**identifications**", and elements of $a \overset{\text{s}}{=} b$ as "**strict equalities**". Note that $a = b$ only eliminates into fibrant types, while $a \overset{\text{s}}{=} b$ eliminates into any type, fibrant or not. Consequently, for any fibrant type A and $a, b : A$, we have a map $(a \overset{\text{s}}{=} b) \to (a = b)$. We sometimes use this implicitly to "coerce" a strict equality to an identification. Similarly, we will frequently prove statements by induction on the strict natural numbers $\mathbb{N}^s$; there, we do not need to pay attention to the return type. We write $A : \mathcal{U}$ to indicate that A is a fibrant, or inner, type.

We write $A \simeq B$ for the type of equivalences between two (necessarily fibrant) types A and B, in the usual sense of HoTT/UF. The *truncation level* of a fibrant type is defined as in [11], with contractible types the (-2)-types, and an n-type being a type whose homotopical identity types are $(n-1)$-types. A *proposition* is a (-1)-type, and a *set* is a 0-type.

An *s*-**category** C (see also [4, Definition 3.1]) is given by the following data ("s" for "strict"):

1. A type C_0 of *objects* (also often denoted C);
2. For each $x, y : C$ a type $C(x, y)$ of *arrows*;
3. For each $x : C$ an arrow $1 : C(x, x)$; and

4. A *composition* map $\circ : C(y, z) \to C(x, y) \to C(x, z)$ that is strictly associative and for which 1 is a strict left and right unit.

A universe $\mathcal{U}^s$ gives rise to an s-category, also called $\mathcal{U}^s$, with objects $A : \mathcal{U}^s$ and morphisms $\mathcal{U}^s(A, B) :\equiv A \to B$. An *s-functor* $F : C \to \mathcal{D}$ consists of a function $F_0 : C_0 \to \mathcal{D}_0$ and functions $F_{x,y} : C(x, y) \to \mathcal{D}(F_0 x, F_0 y)$ preserving identity and composition up to strict equality. We denote both F_0 and $F_{x,y}$ by just F. A *strict natural transformation* $\alpha : F \Rightarrow G : C \to \mathcal{D}$ consists of a family of morphisms $(\alpha_x : \mathcal{D}(Fx, Gx))_{x:C_0}$ satisfying the naturality axiom strictly.

Our signatures will involve strict equality, and will thus live in the outer level of 2LTT. For a fixed signature, the types of structures, of maps between structures, and of in-discernibilities within a structure, will live entirely within the fibrant fragment of 2LTT.

4 Signatures and Structures

In traditional logic, a *signature* specifies the sorts, functions, and relations of a structure. A signature in dependent type theory must also specify the dependencies between sorts; Makkai [9] observed that this enables relations and, to a certain extent, functions, to be expressed merely in terms of sorts. Thus we could adapt Makkai's FOLDS-signatures to 2LTT and define structures as s-functors to $\mathcal{U}$ that are "Reedy fibrant." However, it will be more convenient to formulate the notions of signature and structure *inductively*, leading to a more general class of signatures.

Consider the FOLDS-signature $\mathcal{L}_{rg}$, for which a naïve structure consists of (fibrant) types and families $MO : \mathcal{U}$, $MA : MO \to MO \to \mathcal{U}$, and $MI : \prod_{(x:MO)} MA(x, x) \to \mathcal{U}$. If we strip off the *top* sort I, the resulting structure contains only MO and MA, and an inductive definition can be formulated along these lines. But our inductions will be "bottom-up", so we want to strip off the *bottom* sort O. Once $MO : \mathcal{U}$ is fixed, the rest of an $\mathcal{L}_{rg}$-structure is determined by an ordinary structure over a *derived* signature $(\mathcal{L}_{rg})'_{MO}$, with a rank-0 sort $A(x, y)$ for each $x, y : MO$, and a rank-1 sort $I(x)$ for each $x : MO$, with morphisms $I(x) \to A(x, x)$. That is, we take the "indexing" of all sorts by O and move it "outside" the signature, incorporating it into the types of sorts.[3]

This notion of *derived FOLDS-signature* determines the notion of structure: a structure for $\mathcal{L}$ of height $p > 0$ consists, inductively, of a family $M_\perp : \mathcal{L}(0) \to \mathcal{U}$ and a structure for $\mathcal{L}'_{M_\perp}$ (which is of height $p-1$). We can therefore abstract away from the inverse category underlying a FOLDS-signature, remembering only that each signature $\mathcal{L}$ of height $p > 0$ has (1) a type $\mathcal{L}(0)$, and (2) for any $M_\perp : \mathcal{L}(0) \to \mathcal{U}$, a signature $\mathcal{L}'_{M_\perp}$ of height $p - 1$.

---

[3]This would be impossible if our inverse categories were metatheoretic in the ordinary sense, e.g., syntactic and externally finite. 2LTT is just right.

Definition 4.1 (Abstract signature). We define a family of s-categories $\mathrm{Sig}(n)$ of signatures of height n by induction. Let $\mathrm{Sig}(0)$ be the trivial s-category on $\mathbf{1}$.

An object $\mathcal{L}$ of $\mathrm{Sig}(n + 1)$ consists of

1. a fibrant type $\mathcal{L}_\perp : \mathcal{U}$;
2. a functor $\mathcal{L}' : (\mathcal{L}_\perp \to \mathcal{U}) \to \mathrm{Sig}(n)$, where $\mathcal{L}_\perp \to \mathcal{U}$ is the functor s-category from the discrete s-category $\mathcal{L}_\perp$ to the canonical s-category $\mathcal{U}$.

Arguments of $\mathcal{L}'$ will be written as subscripts, as in $\mathcal{L}'_M$.

For $\mathcal{L}, \mathcal{M} : \mathrm{Sig}(n + 1)$, an element α of $\mathrm{hom}_{\mathrm{Sig}(n+1)}(\mathcal{L}, \mathcal{M})$ consists of the following:

1. a function $\alpha_\perp : \mathcal{L}_\perp \to \mathcal{M}_\perp$
2. a strict natural transformation α' as in the diagram

Arguments of α' will also be written as subscripts, as in α'_M.

Composition and identities are given by function composition and identity at $\perp$, and inductively for the derivative. Similarly, the categorical laws are easily proved by induction.

Similarly, we define $\mathcal{L}$-*structures* inductively for $n : \mathbb{N}^s$ and each $L : \mathrm{Sig}(n)$. The rank-0 part of an $\mathcal{L}$-structure is a type family $M_\perp : \mathcal{L}_\perp \to \mathcal{U}$, while the rest of an $\mathcal{L}$-structure consists of a structure for the derived signature $\mathcal{L}'_{M_\perp}$.

Definition 4.2 ($\mathcal{L}$-structure). If $\mathcal{L} : \mathrm{Sig}(0)$, we define the type of $\mathcal{L}$-*structures* to be $\mathrm{Str}(\mathcal{L}) :\equiv \mathbf{1}$.

If $\mathcal{L} : \mathrm{Sig}(n + 1)$, we define the type of $\mathcal{L}$-*structures* to be

$$\mathrm{Str}(\mathcal{L}) :\equiv \sum_{M_\perp : \mathcal{L}_\perp \to \mathcal{U}} \mathrm{Str}(\mathcal{L}'_{M_\perp}).$$

For $\mathcal{L} : \mathrm{Sig}(n + 1)$ we write $M :\equiv (M_\perp, M') : \mathrm{Str}(\mathcal{L})$.

Lemma 4.3. *For any signature $\mathcal{L}$, the type $\mathrm{Str}(\mathcal{L})$ of $\mathcal{L}$-structures is fibrant.* □

In the rest of the paper we work exclusively with abstract signatures, calling them simply "signatures". However, since most intended examples arise naturally as FOLDS-signatures, we need to be able to translate FOLDS-signatures to abstract ones. To this end, we define general FOLDS-signatures in 2LTT and translate them to abstract signatures, in such a way that "Reedy fibrant" diagrams on a FOLDS-signature coincide with structures for the corresponding abstract signature. In our FOLDS-signatures, the equations of Figure 1 are formulated modulo *strict equality*. This rather technical part of our work is explained in the extended version [3] of this article. Here we sketch the results of the translation for the examples in Figure 1.

Examples 4.4. All three examples have only one sort of rank 0, so that $\mathcal{L}_\perp = \{O\} = \mathbf{1}$, and $M_\perp$ consists of a single type MO. Moreover, since all three examples have only one

sort A of rank 1 that depends on O twice, their derivatives $\mathcal{L}'_{MO}$ have $(\mathcal{L}'_{MO})_\perp = MO \times MO$, and $(M')_\perp$ is a single type family $MA : MO \times MO \to \mathcal{U}$. Finally, since all three have height 3, $(\mathcal{L}'_{MO})'_{MA}$ has height 1, hence is just a single type.

- For $\mathcal{L}_{rg}$, this type is $\sum_{(x:MO)} MA(x, x)$, so that a structure is completed by a type family

$$MI : \left(\sum\nolimits_{(x:MO)} MA(x, x)\right) \to \mathcal{U}.$$

- For $\mathcal{L}_{cat}$, this type is

$$\left(\sum\nolimits_{(x,y,z:MO)} MA(x, y) \times MA(y, z) \times MA(x, z)\right)$$
$$+ \left(\sum\nolimits_{(x:MO)} MA(x, x)\right),$$

so that a structure is completed by a type family MI as above together with

$$MT : \left(\sum\nolimits_{(x,y,z:MO)} MA(x, y) \times MA(y, z) \times MA(x, z)\right) \to \mathcal{U}.$$

- Finally, for $\mathcal{L}_{cat+E}$, this type is

$$\left(\sum\nolimits_{(x,y,z:MO)} MA(x, y) \times MA(y, z) \times MA(x, z)\right)$$
$$+ \left(\sum\nolimits_{(x:MO)} MA(x, x)\right) + \left(\sum\nolimits_{(x,y:MO)} MA(x, y) \times MA(x, y)\right),$$

so that a structure is completed by type families MI and MT as above together with

$$ME : \left(\sum\nolimits_{(x,y:MO)} MA(x, y) \times MA(x, y)\right) \to \mathcal{U}.$$

5 (Iso)morphisms of Structures

The definition of structures for signatures doesn't require the fact that signatures form an s-category. But defining *morphisms* of structures will require the pullback of an $\mathcal{M}$-structure along a morphism $\alpha : \mathcal{L} \to \mathcal{M}$ of signatures, defined as follows.

Definition 5.1. For any $\alpha : \hom_{Sig(n)}(\mathcal{L}, \mathcal{M})$, we define the **pullback** $\alpha^* : Str(\mathcal{M}) \to Str(\mathcal{L})$ inductively as follows.

If $n :\equiv 0$, then let $\alpha^* : Str(\mathcal{M}) \to Str(\mathcal{L})$ be the identity.

If $n > 0$, consider $M : Str(\mathcal{M})$. We let $(\alpha^* M)_\perp$ be $M_\perp \circ \alpha_\perp$. By induction, the morphism

$$\alpha'_{M_\perp} : \hom_{Sig(n-1)}(\mathcal{L}'_{M_\perp \circ \alpha_\perp}, \mathcal{M}'_{M_\perp})$$

produces a $(\alpha'_{M_\perp})^* : Str(\mathcal{M}'_{M_\perp}) \to Str(\mathcal{L}'_{M_\perp \circ \alpha_\perp})$, so we set $(\alpha^* M)' :\equiv (\alpha'_{M_\perp})^* M'$.

Pullback is functorial: pullback along a composition of signature morphisms is the composition of pullbacks, and pullback along an identity morphism is the identity.

We now inductively define *morphisms* between structures of a given signature, making $Str(\mathcal{L})$ into an s-category.

Definition 5.2 (Morphism of structures). Consider $\mathcal{L} : Sig(n)$ and $M, N : Str(\mathcal{L})$.

When $n :\equiv 0$, we let $\hom_{Str(\mathcal{L})}(M, N) :\equiv \mathbf{1}$.

When $n > 0$, a morphism $f : \hom_{Str(\mathcal{L})}(M, N)$ consists of

1. $f_\perp : \prod_{(K:\mathcal{L}_\perp)} M_\perp(K) \to N_\perp(K)$
2. $f' : \hom_{Str(\mathcal{L}'_{M_\perp})}(M', (\mathcal{L}'_{f_\perp})^* N')$.

Lemma 5.3. *For a signature $\mathcal{L}$ and $\mathcal{L}$-structures M and N, the type of morphisms from M to N is fibrant.* $\quad\square$

As a stepping-stone to our SIP for univalent $\mathcal{L}$-structures, we show that *all* $\mathcal{L}$-structures satisfy a tautological "level-wise" form of univalence.

Definition 5.4 (Isomorphism of structures). Consider $\mathcal{L} : Sig(n)$ and $M, N : Str(\mathcal{L})$.

If $n :\equiv 0$, we define every $f : \hom_{Str(\mathcal{L})}(M, N)$ to be an $\mathcal{L}$-isomorphism. That is, we define $isIso_\mathcal{L}(f) :\equiv \mathbf{1}$.

For $n > 0$, $f : \hom_{Str(\mathcal{L})}(M, N)$ is an $\mathcal{L}$-*isomorphism* when

1. $f_\perp(K)$ is an equivalence of types for all $K : \mathcal{L}_\perp$, and
2. f' is an $\mathcal{L}'_{M_\perp}$-isomorphism.

That is, let

$$isIso_\mathcal{L}(f) :\equiv \left(\prod_{(K:L_\perp)} isEquiv(f_\perp(K))\right) \times isIso_{\mathcal{L}'_{M_\perp}}(f').$$

We denote the type of $\mathcal{L}$-isomorphisms between two $\mathcal{L}$-structures M, N by $M \cong_\mathcal{L} N$, or simply $M \cong N$.

Lemma 5.5. *For any morphism $f : M \to N$ between two $\mathcal{L}$-structures, the type $isIso_\mathcal{L}(f)$ is fibrant and a proposition.* $\quad\square$

Definition 5.6 (Identity isomorphism). When $n :\equiv 0$, we define i_M to be the canonical element in $\mathbf{1}$.

Otherwise, we have $1_{\mathcal{L}'_{M_\perp}} \overset{s}{=} \mathcal{L}'_{1_{M_\perp}}$ (from functoriality of $\mathcal{L}'$), hence $u : (1_{\mathcal{L}'_{M_\perp}})^* M' \overset{s}{=} (\mathcal{L}'_{1_{M_\perp}})^* M'$. We also have a strict equality $v : M' \overset{s}{=} (1_{\mathcal{L}'_{M_\perp}})^* M'$ (by the functoriality of pullback). Then we set i_M to be the pair $(1_{M_\perp}, idtoiso(v \cdot u))$.

Now define $idtoiso : \prod_{(M,N:Str(\mathcal{L}))}(M = N) \to (M \cong N)$ by sending $refl_M$ to i_M.

Proposition 5.7. *For structures M, N of a signature $\mathcal{L}$, the canonical map*

$$idtoiso_{M,N} : (M = N) \to (M \cong N)$$

is an equivalence of types.

Proof. When $n :\equiv 0$, $idtoiso : \mathbf{1} \to \mathbf{1}$, hence is an equivalence.

Let $ua : (M_\perp \cong N_\perp) \to (M_\perp = N_\perp)$ be given by the univalence axiom. First we show that $ua(e)^{-1}{}_*(N') = (\mathcal{L}'_e)^* N'$ for any $e : M_\perp \cong N_\perp$, where $ua(e)^{-1}{}_*$ denotes transport along $ua(e)^{-1}$. Now the square in Figure 2 commutes (up to $=$) since both functions $(M_\perp = N_\perp) \to Str(\mathcal{L}'_{N_\perp}) \to Str(\mathcal{L}'_{M_\perp})$ send $refl_{M_\perp}$ to $1_{Str(\mathcal{L}'_M)}$ (by strict functoriality of the pullback). Precomposing these with ua, we find that $(\mathcal{L}'_e)^* N' = ua(e)^{-1}{}_*(N')$. Now we have that

$$(M = N) = \sum_{p:M_\perp = N_\perp} M' = p^{-1}{}_*(N')$$
$$= \sum_{e:M_\perp \cong N_\perp} M' = ua(e)^{-1}{}_*(N')$$
$$= \sum_{e:M_\perp \cong N_\perp} M' = (\mathcal{L}'_e)^* N'$$

$$(M_\perp \cong N_\perp) \xrightarrow{\text{ua}} (M_\perp = N_\perp) \xrightarrow{\text{idtoiso}} (M_\perp \cong N_\perp) \xrightarrow{\mathcal{L}'_-} \hom_{\text{Sig}(n)}(\mathcal{L}'_{M_\perp}, \mathcal{L}'_{N_\perp})$$

with vertical map $(-)^{-1}$ on the left and $(-)^*$ on the right:

$$(N_\perp = M_\perp) \xrightarrow{\quad (-)_* \quad} \text{Str}(\mathcal{L}'_{N_\perp}) \to \text{Str}(\mathcal{L}'_{M_\perp})$$

Figure 2. Diagram for Proof of Proposition 5.7

$$= \sum_{e:M_\perp \cong N_\perp} M' \cong (\mathcal{L}'_e)^* N'$$

$$\equiv (M \cong N)$$

where the second identification is the univalence axiom and the fourth is our inductive hypothesis. This equivalence, from left to right, is idtoiso$_{M,N}$.   □

Proposition 5.7 relies on the univalence axiom; conversely, the univalence axiom can be recovered as an instance of Proposition 5.7, for the signature consisting of just one sort.

Example 5.8. When precategories are regarded as $\mathcal{L}_{\text{cat}}$-structures, their isomorphisms are the *isomorphisms of precategories* from [2, Def. 6.9] and [11, Def. 9.4.8]: functors that induce equivalences on hom-types and also equivalences on types of objects (relative to homotopical *identifications* of objects, not isomorphisms in the category structure).

Remark 5.9. We expect that isomorphisms of structures can equivalently be characterized via the existence of a structure morphism in the other direction and composites that are (homotopically) identical to identities.

The analogue for $\mathcal{L}$-structures of *equivalences* of precategories, called *(split-surjective) equivalences* of $\mathcal{L}$-structures, will be introduced in Section 7. Our main result, Theorem 7.6, will be that between univalent $\mathcal{L}$-structures these are also equivalent to identifications. However, first we have to define univalence of $\mathcal{L}$-structures.

6 Indiscernibility and Univalence

In this section we define indiscernibility of objects within an $\mathcal{L}$-structure. We then define a structure to be univalent when indiscernibility coincides with identification of objects.

Let M be an $\mathcal{L}$-structure, $K : \mathcal{L}_\perp$, and $a, b : M_\perp K$. To define indiscernibilities from a to b, we consider a new $\mathcal{L}$-structure obtained by adding to M one element at sort K: a "joker" element. We can substitute this new element by a or by b; below, we call the obtained structures $\partial_a M$ and $\partial_b M$, respectively. An indiscernibility from a to b will be defined below to be an isomorphism of structures from $\partial_a M$ to $\partial_b M$ that is the identity on all the sorts not depending on the joker element. Intuitively, this means that a and b are isomorphic when one cannot discern one from the other using the rest of the structure M. To make this intuition formal, we need two auxiliary definitions:

Definition 6.1. Consider $L : \mathcal{U}$, $K : L$, $M : L \to \mathcal{U}$, $a : M(K)$. We define the **indicator function of** K to be

$$[K] :\equiv \lambda x. (x = K) : L \to \mathcal{U}$$

and we define the function $a : \prod_{(x:L)}[K](x) \to M(x)$ by sending $\text{refl}_K : [K](K)$ to $a : M(K)$.

Below we consider the pointwise disjoint union $M + [K]$ in $L \to \mathcal{U}$, the canonical injection $\iota_M : \prod_{(x:L)} M(x) \to (M + [K])(x)$, and the induced function $\langle 1_M, a \rangle : \prod_{(x:L)}(M + [K])(x) \to M(x)$.

Definition 6.2. Consider $\mathcal{L} : \text{Sig}(n+1)$, $K : \mathcal{L}_\perp$, $M : \text{Str}(\mathcal{L})$, $a : M_\perp(K)$. Define

$$\partial_a M :\equiv (\mathcal{L}'_{\langle 1_{M_\perp}, a\rangle})^* M' : \text{Str}(\mathcal{L}'_{M_\perp + [K]}).$$

Note that we require the type $\mathcal{L}_\perp$ to be fibrant so that the fibrant indicator function $[K]$ exists.

Now we can define the type of indiscernibilities between objects within an $\mathcal{L}$-structure:

Definition 6.3 (Indiscernibility). Consider $\mathcal{L} : \text{Sig}(n+1)$, $K : \mathcal{L}_\perp$, $M : \text{Str}(\mathcal{L})$, $a, b : M_\perp(K)$. We define the type of **indiscernibilities from a to b** to be

$$(a \asymp b) :\equiv \sum_{p: \, \partial_a M = \partial_b M} \epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_{M_\perp}})^* p \cdot \epsilon_b =_{M'=M'} \text{refl}_{M'},$$

where ϵ_x is the concatenated identification

$$(\mathcal{L}'_{\iota_{M_\perp}})^* \partial_x M' \equiv (\mathcal{L}'_{\iota_{M_\perp}})^* (\mathcal{L}'_{\langle 1_{M_\perp}, x\rangle})^* M'$$
$$= (\mathcal{L}'_{\langle 1_{M_\perp}, x\rangle \circ \iota_{M_\perp}})^* M' = (\mathcal{L}'_{1_{M_\perp}})^* M' = M'.$$

Lemma 6.4. *The type $a \asymp b$ of Definition 6.3 is fibrant.*   □

Remark 6.5. Using identification instead of isomorphism of structures in Definition 6.3 is justified by Proposition 5.7.

Lemma 6.6. *The type of indiscernibilities $a \asymp b$ of Definition 6.3 is equivalent to the type*

$$\sum_{p: \, \partial_a M = \partial_b M} (\mathcal{L}'_{\iota_{M_\perp}})^* p = \epsilon_a \cdot \epsilon_b^{-1}. \qquad\qquad □$$

We now define *univalence of $\mathcal{L}$-structures*. For this, we first need to define the canonical map from identifications to indiscernibilities.

Definition 6.7 (Identity indiscernibility). For $\mathcal{L} : \text{Sig}(n+1)$, $K : \mathcal{L}_\perp$, $M : \text{Str}(\mathcal{L})$, and $m : M_\perp(K)$, we define the indiscernibility $1 : m \asymp m$. Let $M : \text{Str}(\mathcal{L})$. For any $a : M_\perp(K)$, we have $\text{refl}_{\partial_a M} : \partial_a M = \partial_a M$. Then

$$\epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_M})^* (\text{refl}_{\partial_a M}) \cdot \epsilon_a$$

$$\overset{s}{=} \epsilon_a^{-1} \cdot \mathsf{refl}_{(\mathcal{L}'_{\iota_M})^* \partial_a M} \cdot \epsilon_a$$
$$= \mathsf{refl}_{M'} \ ,$$

where the second identification uses the groupoidal properties of types. This gives the desired indiscernibility.

Definition 6.8. Consider $\mathcal{L} : \mathsf{Sig}(n+1), K : \mathcal{L}_\perp, M : \mathsf{Str}(\mathcal{L})$. For any $a, b : M_\perp(K)$, let $\mathsf{idtoiso}_{a,b} : (a = b) \to (a \asymp b)$ be the function which sends refl_a to the identity indiscernibility exhibited in Definition 6.7.

We say that M is **univalent at** K if for all $a, b : M_\perp(K)$, $\mathsf{idtoiso}_{a,b} : (a = b) \to (a \asymp b)$ is an equivalence.

Definition 6.9 (Univalence of structures). We define by induction what it means for a structure of a signature $\mathcal{L} : \mathsf{Sig}(n)$ to be univalent.

When $n :\equiv 0$, every structure $M : \mathsf{Str}(\mathcal{L})$ is univalent.

Otherwise, a structure $M : \mathsf{Str}(\mathcal{L})$ is univalent if M is univalent at all $K : \mathcal{L}_\perp$ and M' is univalent.

Let $\mathsf{uStr}(\mathcal{L})$ denote the type of univalent structures of $\mathcal{L}$.

Lemma 6.10. *Let $\mathcal{L}$ be a signature.*
- *The type $\mathsf{uStr}(\mathcal{L})$ is fibrant.*
- *For any $\mathcal{L}$-structure, "being univalent" is a proposition.*
- *Identification of univalent $\mathcal{L}$-structures corresponds to identification of the underlying $\mathcal{L}$-structures.* □

Example 6.11. Suppose $\mathcal{L}$ has height 1, hence is just a type $\mathcal{L}_\perp$. Consider an $\mathcal{L}$-structure $M : \mathcal{L}_\perp \to \mathcal{U}$ and $a, b : M(K)$. Then $\partial_a M$ and $\partial_b M$ are structures for the trivial signature of height 0, hence uniquely identified; thus $(a \asymp b) = \mathbf{1}$. So any structure of a signature $\mathcal{L}$ of height 1 is univalent just when it consists entirely of propositions.

Example 6.12. Recall from Examples 4.4 that for $\mathcal{L} = \mathcal{L}_{\mathrm{cat+E}}$, we have $\mathcal{L}_\perp = 1$, $M_\perp = MO : \mathcal{U}$, $\mathcal{L}'_{MO\perp} = MO \times MO$, and $(M')_\perp = MA : MO \times MO \to \mathcal{U}$, while M'' consists of the sorts $MT_{x,y,z}(f, g, h)$, $MI_x(f)$, and $ME_{x,y}(f, g)$. By Example 6.11, M'' is univalent just when all these types are propositions. Now for any $a, b : MO$, we have

$$(MA + [A(a, b)])(x, y) = MA(x, y) + ((a = x) \times (b = y)).$$

Thus, the height-1 signature $(\mathcal{L}'_{MO})'_{MA+[A(a,b)]}$ is

$$\left(\sum\nolimits_{(x,y,z:MO)} (MA(x, y) + ((a = x) \times (b = y))) \right.$$
$$\times (MA(y, z) + ((a = y) \times (b = z)))$$
$$\times (MA(x, z) + ((a = x) \times (b = z))))$$
$$+ \left(\sum\nolimits_{(x:MO)} (MA(x, x) + ((a = x) \times (b = x))) \right)$$
$$+ \left(\sum\nolimits_{(x,y:MO)} (MA(x, y) + ((a = x) \times (b = y))) \right.$$
$$\left. \times (MA(x, y) + ((a = x) \times (b = y))) \right).$$

By distributing $\sum$ and $\times$ over $+$ and contracting some singletons, this is equivalent to

$$\left(\sum\nolimits_{(x,y,z:MO)} MA(x, y) \times MA(y, z) \times MA(x, z) \right) \tag{31}$$
$$+ \left(\sum\nolimits_{(z:MO)} MA(b, z) \times MA(a, z) \right) \tag{32}$$
$$+ \left(\sum\nolimits_{(x:MO)} MA(x, a) \times MA(x, b) \right) \tag{33}$$
$$+ \left(\sum\nolimits_{(y:MO)} MA(a, y) \times MA(y, b) \right) \tag{34}$$
$$+ \left((a = b) \times MA(a, b) \right) \tag{35}$$
$$+ \left(MA(a, a) \times (b = b) \right) \tag{36}$$
$$+ \left((a = a) \times MA(b, b) \right) \tag{37}$$
$$+ \left((a = b) \times (a = a) \times (b = b) \right) \tag{38}$$
$$+ \left(\sum\nolimits_{(x:MO)} MA(x, x) \right) \tag{39}$$
$$+ \left((a = b) \right) \tag{40}$$
$$+ \left(\sum\nolimits_{(x,y:MO)} MA(x, y) \times MA(x, y) \right) \tag{41}$$
$$+ \left(MA(a, b) \right) \tag{42}$$
$$+ \left(MA(a, b) \right) \tag{43}$$
$$+ \left((a = a) \times (b = b) \right). \tag{44}$$

Thus for $f, g : MA(a, b)$, an identification $\partial_f M = \partial_g M$ consists of equivalences between instances of the predicates MT, MI, ME indexed over the types (31)–(44). The condition on restriction along ι says that the equivalences corresponding to (31), (39), and (41) are the identity, while those corresponding to (32)–(34), (35)–(38), (40), and (42)–(44) yield respectively the equivalences (20)–(22), (23)–(26), (27), and (28)–(30) from Section 2.3. Hence, indiscernibilities $f \asymp g$ in the sense of Definition 6.3 coincide with the indiscernibilities from Definition 2.10.

Now moving back down to the bottom rank, an $(\mathcal{L}'_{MO})$-structure consists of $MA : MO \times MO \to \mathcal{U}$ together with appropriately typed families MT, MI, and ME. Since $(MO + [O]) = MO + \mathbf{1}$, for $a : MO$ the 0th rank of $\partial_a M$ is

$$(\partial_a M) A : (MO + \mathbf{1}) \times (MO + \mathbf{1}) \to \mathcal{U}$$

or equivalently

$$(\partial_a M) A : (MO \times MO) + MO + MO + \mathbf{1} \to \mathcal{U}$$

consisting of the types $(MA(x, y))_{x,y:MO}$, $(MA(a, y))_{y:MO}$, $(MA(x, a))_{x:MO}$, and $MA(a, a)$. The 1st rank consists of MT, MI, and ME pulled back appropriately to these families. Thus, an identification $\partial_a M = \partial_b M$ consists of equivalences

$$MA(x, y) \simeq MA(x, y) \tag{45}$$
$$MA(x, a) \simeq MA(x, b) \tag{46}$$
$$MA(a, y) \simeq MA(b, y) \tag{47}$$
$$MA(a, a) \simeq MA(b, b) \tag{48}$$

for all $x, y : MO$ that respect the predicates MT, MI, ME. The condition on restriction along ι says that the equivalences (45) are the identity, while the remaining (46)–(48) correspond respectively to the equivalences $\phi_{x\bullet}$, $\phi_{\bullet y}$, and $\phi_{\bullet\bullet}$ from Section 2.3. Finally, respect for MT, MI, ME specializes to (9)–(16) together with analogous equivalences for E that are trivial under "standardness" of identifications. Thus, indiscernibilities $a \asymp b$ in the sense of Definition 6.3 coincide with the indiscernibilities from Definition 2.5.

Our first general observations about univalent structures give truncation bounds for their sorts and for the type of such structures.

Theorem 6.13. *Let* $\mathcal{L} : \mathrm{Sig}(n + 1)$, $M : \mathrm{uStr}(\mathcal{L})$, $K : \mathcal{L}_\perp$. *Then* $M_\perp(K)$ *is an* $(n - 1)$*-type.*

Theorem 6.14. *Let* $\mathcal{L} : \mathrm{Sig}(n)$. *The type of univalent* $\mathcal{L}$*-structures is an* $(n - 1)$*-type.*

Proof of Theorems 6.13 and 6.14. Define the following types.

$$P(n) :\equiv \prod_{\mathcal{L}:\mathrm{Sig}(n+1)} \prod_{M:\mathrm{uStr}(\mathcal{L})} \prod_{K:\mathcal{L}_\perp} \mathrm{is}\text{-}(n-1)\text{-type}(M_\perp(K))$$

$$Q(n) :\equiv \prod_{\substack{\mathcal{M},\mathcal{N}:\mathrm{Sig}(n) \\ \alpha:\mathrm{hom}(\mathcal{M},\mathcal{N})}} \prod_{N:\mathrm{uStr}(\mathcal{N})} \mathrm{is}\text{-}(n-2)\text{-type}(\alpha^*N = \alpha^*N)$$

The type $P(n)$ is the statement of Theorem 6.13, and the type $Q(n)$ implies the statement of Theorem 6.14 by [11, Thm. 7.2.7]. We prove $P(n)$ and $Q(n)$ simultaneously.

For $P(n)$, we need to show that $a =_{M_\perp K} b$ is an $(n-2)$-type for all $\mathcal{L} : \mathrm{Sig}(n + 1), M : \mathrm{uStr}(\mathcal{L}), K : \mathcal{L}_\perp, a, b : M_\perp K$. But since M is univalent, this type is equivalent to

$$(a \asymp b) \equiv \sum_{e:\partial_a M = \partial_b M} \epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_M})^* p \cdot \epsilon_b =_{M'=M'} \mathrm{refl}_{M'}.$$

Thus, it will suffice to show that $\partial_a M = \partial_b M$ and $\epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_M})^* p \cdot \epsilon_b =_{M'=M'} \mathrm{refl}_{M'}$ are $(n-2)$-types.

To show $P(0)$ and $Q(0)$ consider $\mathcal{L} : \mathrm{Sig}(1), M : \mathrm{uStr}(\mathcal{L}), K : \mathcal{L}_\perp, a, b : M_\perp K, \mathcal{M}, \mathcal{N} : \mathrm{Sig}(0), \alpha : \mathrm{hom}(\mathcal{M}, \mathcal{N}), N : \mathrm{uStr}(\mathcal{N})$. We have that $M', \partial_a M, \partial_b M, \alpha^*N : \mathbf{1}$ so the types $\partial_a M = \partial_b M$, $\epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_M})^* p \cdot \epsilon_b =_{M'=M'} \mathrm{refl}_{M'}$, and $\alpha^*N = \alpha^*N$ are contractible. Thus, $P(0)$ and $Q(0)$ hold.

Suppose that $P(n)$ and $Q(n)$ hold. We first show $Q(n + 1)$. Consider $\mathcal{M}, \mathcal{N} : \mathrm{Sig}(n + 1), \alpha : \mathrm{hom}(\mathcal{M}, \mathcal{N}), N : \mathrm{uStr}(\mathcal{N})$. We have that

$$(\alpha^*N = \alpha^*N) \simeq \Sigma_{e:(\alpha^*N)_\perp=(\alpha^*N)_\perp}(\alpha^*N)' = e_*(\alpha^*N)'$$
$$\equiv \Sigma_{e:(N_\perp \circ \alpha_\perp)=(N_\perp \circ \alpha_\perp)}(\alpha'_{N_\perp})^*N' = e_*(\alpha'_{N_\perp})^*N'.$$

Our inductive hypothesis $Q(n)$ ensures that $(\alpha'_{N_\perp})^*N' = (\alpha'_{N_\perp})^*N'$ is an $(n-2)$-type, and hence $(\alpha'_{N_\perp})^*N' = e_*(\alpha'_{N_\perp})^*N'$ is an $(n-1)$-type by [11, Thm. 7.2.7]. It remains to show that $(N_\perp \circ \alpha_\perp) = (N_\perp \circ \alpha_\perp)$ is an $(n-1)$-type. Note that N is a univalent structure of an $(n + 1)$-signature, and our inductive hypothesis $P(n)$ then implies that for all $K : \mathcal{N}_\perp$, the type $N_\perp(K)$ is an $(n - 1)$-type. Then since $(N_\perp \circ \alpha_\perp)$ is a function which takes values in $(n - 1)$-types, we can conclude that $(N_\perp \circ \alpha_\perp) = (N_\perp \circ \alpha_\perp)$ is an $(n - 1)$-type [11, Thm. 7.1.9]. Thus, $Q(n + 1)$ holds.

To show that $P(n + 1)$ holds, consider $\mathcal{L} : \mathrm{Sig}(n + 2), M : \mathrm{uStr}(\mathcal{L}), K : \mathcal{L}_\perp, a, b : M_\perp K$. By [11, Thm. 7.2.7], $Q(n + 1)$ implies that $\partial_a M = \partial_b M$ and $\epsilon_a^{-1} \cdot (\mathcal{L}'_{\iota_M})^* p \cdot \epsilon_b =_{M'=M'} \mathrm{refl}_{M'}$ are $(n - 2)$-types. Therefore, $P(n + 1)$ holds. □

Example 6.15. For the signature $\mathcal{L}_{\mathrm{cat+E}}$ of height 3, Theorem 6.13 states that the type of objects of a univalent $\mathcal{L}_{\mathrm{cat+E}}$-structure, and hence also of a univalent FOLDS-category, is a 1-type. Theorem 6.14 states that the type of univalent $\mathcal{L}_{\mathrm{cat+E}}$-structures, and hence also the type of univalent FOLDS-categories (as a subtype of the former), is a 2-type.

7 Equivalence of Structures and Higher SIP

Our Higher Structure Identity Principle (Theorem 7.6) requires a notion of equivalence *of* structures that is *a priori* weaker than Definition 5.4. Makkai defined a morphism of structures $\mu : M \to N$ to be *very surjective*[4] if it is "locally" surjective at all sorts, which in our notation means that $\mu_\perp(K)$ is surjective for all K and μ' is (inductively) very surjective. For instance, a morphism between FOLDS-categories is very surjective if (1) it is surjective on objects, (2) it is *full*, i.e., each map $MA(x, y) \to NA(fx, fy)$ is surjective, and (3) it reflects the relations T, I, E. When equality is standard, (3) is equivalent to faithfulness, so this is just a surjective weak equivalence of categories. With this motivation, we use similar terminology for structures over arbitrary signatures.

Definition 7.1 (Split-surjective equivalence). Suppose $f : \mathrm{hom}_{\mathrm{Str}(\mathcal{L})}(M, N)$, where $M, N : \mathrm{Str}(\mathcal{L})$ and $\mathcal{L} : \mathrm{Sig}(n)$. If $n :\equiv 0$, then f is a split-surjective equivalence. For $n > 0$, f is a split-surjective equivalence if

1. $f_\perp(K)$ is a split surjection for every $K : \mathcal{L}_\perp$, and
2. f' is a split-surjective equivalence.

Surjective weak equivalences are defined similarly, but only requiring each $f_\perp(K)$ to be surjective.

Unfortunately, we are currently unable to prove our desired general result with surjective weak equivalences, so for the present we restrict to the split-surjective equivalences. We write $\mathrm{SSEquiv}(f)$ for the type "f is a split-surjective equivalence", which in the inductive case is

$$\mathrm{SSEquiv}(f) :\equiv \left(\prod_{K:\mathcal{L}_\perp} \sum_s (f_\perp K \circ s = 1) \right) \times \mathrm{SSEquiv}(f'),$$

and $(M \twoheadrightarrow N) :\equiv \sum_{(f:\mathrm{hom}_{\mathrm{Str}(\mathcal{L})}(M,N))} \mathrm{SSEquiv}(f)$ for the type of split-surjective equivalences.

Lemma 7.2. *The type* $M \twoheadrightarrow N$ *of split-surjective equivalences is fibrant.* □

Definition 7.3 (From isomorphisms to split-surjective equivalences). Let $f : \mathrm{hom}_{\mathrm{Str}(\mathcal{L})}(M, N)$; we define $U_f : \mathrm{isIso}(f) \to \mathrm{SSEquiv}(f)$ by induction on n. If $n :\equiv 0$, U_f is the identity function on $\mathbf{1}$. For $n > 0$, we use that any equivalence of types is a split surjection, and the inductive hypothesis. Let $\mathrm{isotosse}_{M,N} :\equiv (1, U) : (M \cong N) \to (M \twoheadrightarrow N)$.

---
[4]Very surjective morphisms are also known as "Reedy surjections" and "trivial fibrations".

Definition 7.4 (From identifications to split-surjective equivalences). For $\mathcal{L} : \mathrm{Sig}(n)$ and $M, N : \mathrm{Str}(\mathcal{L})$ we define

$$\mathrm{idtosse} :\equiv \mathrm{isotosse} \circ \mathrm{idtoiso} : (M = N) \to (M \twoheadrightarrow N).$$

Our HSIP states that if M is univalent, then $\mathrm{idtosse}_{M,N}$ is an equivalence. It uses the following lemma.

Lemma 7.5. *Let* $\mathcal{L} : \mathrm{Sig}(n+1)$*,* $M, N : \mathrm{Str}(\mathcal{L})$*,* $f_\perp : M_\perp \to N_\perp$*, and* $e : M' = (\mathcal{L}'_{f_\perp})^* N'$*. Then for* $x, y : M_\perp(K)$*, an indiscernibility* $f_\perp x \asymp f_\perp y$ *produces an indiscernibility* $x \asymp y$*.*

Proof. By path induction on e, we may assume $M' \equiv (\mathcal{L}'_{f_\perp})^* N'$.

Consider the following diagram whose cells commute up to $\overset{s}{=}$ or $=$, as pictured.

$$
\begin{array}{ccccc}
 & & 1_{M_\perp} & & \\
 & & \overset{s}{=} & & \\
M_\perp & \xrightarrow{\iota_{M_\perp}} & M_\perp + [K] & \xrightarrow{\langle 1, x \rangle} & M_\perp \\
\downarrow{f_\perp} & \overset{s}{=} & \downarrow{f_\perp + 1} & = & \downarrow{f_\perp} \\
N_\perp & \xrightarrow{\iota_{N_\perp}} & N_\perp + [K] & \xrightarrow{\langle 1, f_\perp(K)x \rangle} & N_\perp \\
 & & \overset{s}{=} & & \\
 & & 1_{N_\perp} & &
\end{array}
\qquad (49)
$$

This diagram commutes 2-dimensionally, which is to say that the "pasting" of all four displayed identities is strictly equal to the strict (indeed, judgemental) equality $f_\perp \circ 1_{M_\perp} \overset{s}{=} 1_{N_\perp} \circ f_\perp$. Applying the composite s-functor $\mathrm{Str}(\mathcal{L}'_-)$, we obtain:

$$
\begin{array}{ccccc}
 & & 1_{\mathrm{Str}(\mathcal{L}'_{M_\perp})} & & \\
 & & \overset{s}{=}(\epsilon_x) & & \\
\mathrm{Str}(\mathcal{L}'_{M_\perp}) & \xleftarrow{(\mathcal{L}'_{\iota_{M_\perp}})^*} & \mathrm{Str}(\mathcal{L}'_{M_\perp + [K]}) & \xleftarrow{(\mathcal{L}'_{\langle 1, x \rangle})^*} & \mathrm{Str}(\mathcal{L}'_{M_\perp}) \\
\uparrow{(\mathcal{L}'_{f_\perp})^*} & \overset{s}{=}(\alpha) & \uparrow{(\mathcal{L}'_{f_\perp + 1})^*} & =(\beta_x) & \uparrow{(\mathcal{L}'_{f_\perp})^*} \\
\mathrm{Str}(\mathcal{L}'_{N_\perp}) & \xleftarrow{(\mathcal{L}'_{\iota_{N_\perp}})^*} & \mathrm{Str}(\mathcal{L}'_{N_\perp + [K]}) & \xleftarrow{(\mathcal{L}'_{\langle 1, f_\perp x \rangle})^*} & \mathrm{Str}(\mathcal{L}'_{N_\perp}) \\
 & & \overset{s}{=}(\epsilon_{f_\perp x}) & & \\
 & & 1_{\mathrm{Str}(\mathcal{L}'_{N_\perp})} & &
\end{array}
$$

$$(50)$$

which commutes in the same way. Moreover, the upper and lower strict equalities in this diagram are ϵ_x and $\epsilon_{f_\perp x}$ respectively; we call the others α and β_x.

We have an analogous diagram for y, in which the left-hand square α is the same.

Then since $\partial_{f_\perp(K)x} N \equiv (\mathcal{L}'_{\langle 1, f_\perp(K)x \rangle})^* N'$, $M' \equiv (\mathcal{L}'_{f_\perp})^* N'$, and $\partial_x M \equiv (\mathcal{L}'_{\langle 1, x \rangle})^* M'$, we have an identification

$$\beta_x N : (\mathcal{L}'_{f_\perp + 1})^* \partial_{f_\perp(K)x} N = \partial_x M.$$

The same can be shown for y.

Consider an indiscernibility $f_\perp x \asymp f_\perp y$ which consists, by Lemma 6.6, of (1) an identification $i : \partial_{f_\perp x} N = \partial_{f_\perp y} N$ and (2) an identification j between $(\mathcal{L}'_{\iota_{N_\perp}})^* i$ and the concatenation

$$(\mathcal{L}'_{\iota_{N_\perp}})^* (\mathcal{L}'_{\langle 1, f_\perp x \rangle})^* N' \overset{\epsilon_{f_\perp x}}{=} N' \overset{\epsilon_{f_\perp y}^{-1}}{=} (\mathcal{L}'_{\iota_{N_\perp}})^* (\mathcal{L}'_{\langle 1, f_\perp y \rangle})^* N'$$

(which is strict, though i is not).

We need to construct an indiscernibility $x \asymp y$ which consists of (1) an identification $k : \partial_x M = \partial_y M$ and (2) an identification $(\mathcal{L}'_{\iota_{M_\perp}})^* k = \epsilon_x \cdot \epsilon_y^{-1}$.

The first component, k, of our desired indiscernibility $x \asymp y$ is the following concatenation:

$$(\mathcal{L}'_{\langle 1, x \rangle})^* (\mathcal{L}'_{f_\perp})^* N' \overset{\beta_x}{=} (\mathcal{L}'_{f_\perp + 1})^* (\mathcal{L}'_{\langle 1, f_\perp x \rangle})^* N'$$

$$\overset{(\mathcal{L}'_{f_\perp + 1})^* i}{=} (\mathcal{L}'_{f_\perp + 1})^* (\mathcal{L}'_{\langle 1, f_\perp y \rangle})^* N'$$

$$\overset{\beta_y^{-1}}{=} (\mathcal{L}'_{\langle 1, y \rangle})^* (\mathcal{L}'_{f_\perp})^* N'$$

Now we need $(\mathcal{L}'_{\iota_{M_\perp}})^* k = \epsilon_x \cdot \epsilon_y^{-1}$. Consider the commutative diagram in Figure 3 where straight lines denote strict equalities, squiggly lines denote identifications, and double (squiggly) lines denote identifications between identifications. The 2-dimensional identification labeled ν arises from naturality, while those labeled σ arise from the 2-dimensional commutativity of Diag. (50). The concatenation of the three top horizontal identifications in Figure 3 is $(\mathcal{L}'_{\iota_{M_\perp}})^* k$. Thus, Figure 3 exhibits an identification of this with $\epsilon_x \cdot \epsilon_y^{-1}$. □

Theorem 7.6 (Higher SIP). *Consider* $\mathcal{L} : \mathrm{Sig}(n)$ *and* $M, N : \mathrm{Str}(\mathcal{L})$ *such that* M *is univalent. The morphism* $\mathrm{idtosse} : (M = N) \to (M \twoheadrightarrow N)$ *is an equivalence.*

Proof. It suffices to show that each U_f of Definition 7.3 is an equivalence. We proceed by induction on n. When $n :\equiv 0$, each U_f is a endofunction on $\mathbf{1}$, and so is an equivalence.

When $n > 0$, we first construct a map $F_f : \mathrm{SSEquiv}(f) \to \mathrm{isIso}(f)$. Consider an element of $\mathrm{SSEquiv}(f)$: a right inverse $s(K)$ of $f_\perp(K)$ for each $K : \mathcal{L}_\perp$, and $s' : \mathrm{SSEquiv}(f')$. Since M' is univalent, the inductive hypothesis for s' implies f' is an isomorphism; thus it remains to show each $f_\perp(K)$ is an equivalence.

Since $s(K)$ is a right inverse of $f_\perp(K)$, it remains to show that we have $s(K)f_\perp(K)m = m$ for any $m : M_\perp(K)$. We have $f_\perp(K)s(K)f_\perp(K)m = f_\perp(K)m$ and thus $f_\perp(K)s(K)f_\perp(K)m \asymp f_\perp(K)m$. We have already shown that f' is an isomorphism $M' \cong (f_\perp)^* N'$, so by Proposition 5.7, we get $M' = (\mathcal{L}'_{f_\perp})^* N'$. Thus, by Lemma 7.5, we have $s(K)f_\perp(K)m \asymp m$; and since M is univalent this yields $s(K)f_\perp(K)m = m$.

Thus, given our $(\lambda K.s(K), s') : \mathrm{SSEquiv}(f)$, we have constructed an element of $\mathrm{isIso}(f)$; this defines $F_f : \mathrm{SSEquiv}(f) \to \mathrm{isIso}(f)$. Since $\mathrm{isIso}(f)$ is a proposition (by Lemma 5.5), $F_f U_f = 1$. Moreover, we constructed F_f and U_f such that $U_f F_f = 1$.[5] Hence, $U_f : \mathrm{isIso}(f) \to \mathrm{SSEquiv}(f)$ is an equivalence.

Thus, the function $\mathrm{isotosse}_{M,N} : (M \cong_L N) \to (M \twoheadrightarrow N)$ is also an equivalence. Using Proposition 5.7, we find then that $\mathrm{idtosse} : (M = N) \to (M \twoheadrightarrow N)$ is an equivalence. □

---

[5]Since we showed that $f_\perp(K)$ was an equivalence by making $s(K)$ a homotopy inverse of it, and U_f remembers not just the inverse map but one of the homotopies, we technically have to use here the fact that a homotopy inverse of a function g can be enhanced to an element of $\mathrm{isEquiv}(g)$ while changing at most one of the constituent homotopies.
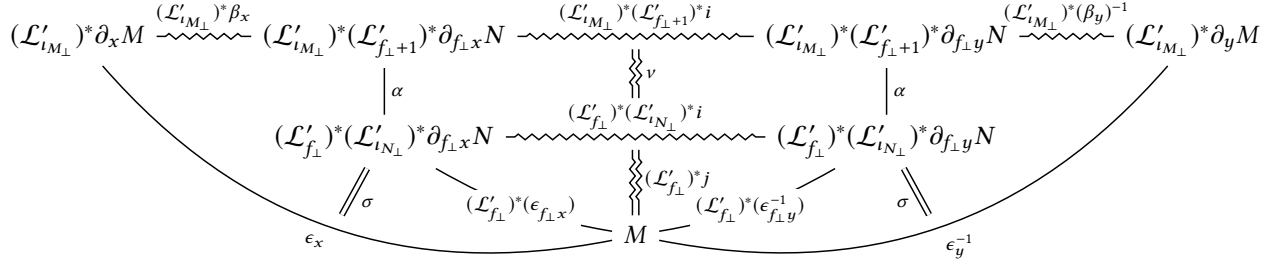
Figure 3. Diagram for proof of Lemma 7.5

Example 7.7. A split-surjective equivalence between univalent FOLDS-categories is the same as a fully faithful and split (essentially) surjective functor, which by [2, Lemma 6.6] is the same as an equivalence of categories. Thus, Theorem 7.6 specializes to [2, Theorem 6.17].

Remark 7.8. With some more care, it is also possible to define a general analogue of non-surjective equivalences of categories and prove an HSIP for them. The idea is to replace surjectivity by essential surjectivity, i.e., surjectivity up to an indiscernibility. However, this is a bit tricky to make precise, because we want these indiscernibilities at all dimensions to live in the codomain structure N and its derivatives; whereas with a naïve inductive definition these indiscernibilities would end up living in pullbacks of these structures such as f^*N' (the codomain of f'), which may not be univalent even if N is. The details can be found in the extended version [3] of this article.

We do not know any analogue of [2, Lemma 6.8], i.e., an HSIP for *weak* equivalences.

8 Axioms and Theories

One of Makkai's goals was to define, for a given (FOLDS)-signature $\mathcal{L}$, a language for properties that are invariant under $\mathcal{L}$-equivalence. Working in 2LTT, we do not need to devise a language for invariant properties ourselves; instead, we can rely on the homotopical fragment of 2LTT to sufficiently constrain our language.[6]

Definition 8.1. Let $\mathcal{L}$ be a signature. An $\mathcal{L}$-**axiom** is a function $\mathrm{Str}(\mathcal{L}) \to \mathrm{Prop}_{\mathcal{U}}$.

Example 8.2. The axioms given in Equations (4) and (5) straightforwardly give rise to axioms for the signature $\mathcal{L}_{\mathrm{cat+E}}$.

Any $\mathcal{L}$-axiom restricts to a predicate on univalent $\mathcal{L}$-structures. Our main result implies that any $\mathcal{L}$-axiom is invariant under equivalence of univalent $\mathcal{L}$-structures:

Corollary 8.3. *Given an $\mathcal{L}$-axiom t, univalent $\mathcal{L}$-structures M, N, and a split-surjective equivalence $M \twoheadrightarrow N$, we have $t(M) \leftrightarrow t(N)$.*

---

[6]An internally-syntactic description of such "axioms" might nevertheless prove useful in the future.

A **theory** is a pair $(\mathcal{L}, T)$ of a signature $\mathcal{L}$ and a family T of $\mathcal{L}$-axioms whose indexing type is fibrant; e.g., a list of five axioms can be specified by a family indexed by the standard finite fibrant type of five elements. A **(univalent) model of a theory** $(\mathcal{L}, T)$ then consists of a (univalent) $\mathcal{L}$-structure M together with a proof $t(M)$ for each axiom t of T. The type of such models is fibrant. In the next section, we discuss some particular theories and their univalent models.

9 Examples of Theories

In this section, we discuss some examples of theories and their (univalent) models.

Example 9.1 (First-order logic). Consider a many-sorted first-order theory with relations and equality (assumed to be a congruence):
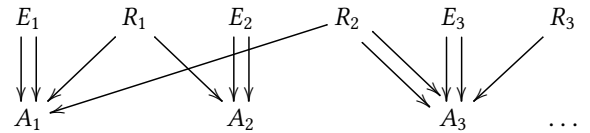
- Univalence at E_i and R_i makes them proposition-valued.
- Univalence at A_i makes it a set whose equality is E_i.

We recover first-order logic with equality. Any instance of this example, with sorts $(A_i)_{i:I}$, is also an instance of the SIP [11, Section 9.9] over Set^I, including the examples of posets, monoids, groups, and fields mentioned in Section 1.2.

Example 9.2 (†-categories). A †-category is a category with coherent isomorphisms $(\_)^{\dagger} : \hom(x, y) \to \hom(y, x)$. A signature for †-categories is as follows,
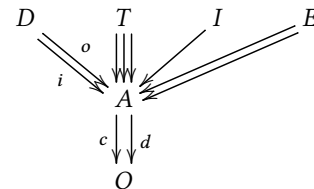
with $co \overset{s}{=} di$ and $do \overset{s}{=} ci$ and the strict equalities of Figure 1. Then for †-categories regarded as structures:

- An indiscernibility between $x, y : MO$ is a *unitary isomorphism* $f : x \cong y$, i.e., one satisfying $f^{-1} = f^{\dagger}$.

- In a univalent †-category, O is the groupoid of objects and unitary isomorphisms.
- Split-surjective equivalences are †-*equivalences*, involving unitary natural isomorphisms.

Further examples of higher structures that can be specified via a suitable signature are:

- a category with a presheaf on it;
- two categories and a(n ana)functor between them;
- two categories and a profunctor between them;
- a displayed category (or fibration) over some base.

These examples are explained in detail in the extended version [3] of this article.

Example 9.3 (T_0-spaces). We end by sketching some signatures whose structures include topological spaces, showing that our abstract signatures are more general than FOLDS-signatures. Since a topology is a structure on one underlying set, it suffices to consider height-2 signatures with $\mathcal{L}_\perp :\equiv \mathbf{1}$, with $\mathcal{L}' : \mathcal{U} \to \mathcal{U}$ remaining to be specified.

A first guess might be $\mathcal{L}'_M :\equiv (M \to \mathrm{Prop}_\mathcal{U})$, so that an $\mathcal{L}$-structure would be a type M with a predicate on its "type of subsets" $M \to \mathrm{Prop}_\mathcal{U}$ representing "is open". Unfortunately, this is not a covariant s-functor. We can make it covariant via direct images (using propositional truncation), but this is not *strictly* functorial, and moreover the resulting morphisms of structures would be open maps rather than continuous ones.

Covariant strict functoriality does hold, however, for the double-powerset functor $M \mapsto ((M \to \mathrm{Prop}_\mathcal{U}) \to \mathrm{Prop}_\mathcal{U})$, so we can use a definition of topological spaces that refers to sets of subsets instead of individual subsets. For instance, a topology is equivalent to a *convergence* relation between filters (which are sets of subsets) and points, hence can be regarded as a particular $\mathcal{L}$-structure with

$$\mathcal{L}'_M :\equiv ((M \to \mathrm{Prop}_\mathcal{U}) \to \mathrm{Prop}_\mathcal{U}) \times M.$$

The covariant functoriality specializes to the direct image of filters, so the $\mathcal{L}$-structure morphisms between topological spaces will be functions that preserve convergence, which is equivalent to continuity. Finally, univalence means that convergence is a proposition, that M is a set, and that two points are identified if exactly the same filters converge to them; the latter is an equivalent way of saying the space is T_0. Of course, not every $\mathcal{L}$-structure is a topological space; we could hope to single out the spaces with a theory in an appropriate logic based on our signatures.

We can also take $\mathcal{L}'_M :\equiv ((M \to \mathrm{Prop}_\mathcal{U}) \to \mathrm{Prop}_\mathcal{U})$ and associate a topological space M to the set of all sets of subsets $\mathcal{T}$ such that all open subsets belong to $\mathcal{T}$. Once again the structure morphisms between topological spaces are the continuous maps, and the univalent such structures are the T_0-spaces. Other topological structures such as uniform spaces and proximity spaces can similarly be represented as structures over suitable height-2 signatures.

10 Conclusion

Using a relativized form of the *identity of indiscernibles*, we defined a general notion of indiscernibility of objects in a categorical structure, yielding a notion of univalence for such structures. These notions depend only on the shape of the structures as specified by the signature, not on any axioms they satisfy. We then showed a Structure Identity Principle for univalent structures that specializes to known results for first-order logic and univalent 1-categories, as well as many other important examples.

Regarding the setting we have chosen for our work, it seems impossible to define a fully coherent notion of signature without 2LTT. A sufficiently-coherent "wild" notion might suffice for our particular results, but further development of the theory may require the fully coherent version. In addition, 2LTT is necessary for treating FOLDS-signatures of arbitrary height (cf. [3]).

Some goals for further work include:

- Removing the splitness condition from Theorem 7.6, as discussed at the end of Section 7.
- Developing a completion operation for structures, i.e., a universal way to turn a structure into a univalent one, generalizing the Rezk completion for categories [2, Section 8].
- Formalizing the results presented here in a computer proof assistant implementing 2LTT.

References

[1] Peter Aczel. 2011. On Voevodsky's Univalence Axiom. In *Mathematical Logic: Proof Theory, Constructive Mathematics*, Samuel R. Buss, Ulrich Kohlenbach, and Michael Rathjen (Eds.). Mathematisches Forschungsinstitut Oberwolfach, Oberwolfach, Chapter 1, 2967. https://doi.org/10.4171/OWR/2011/52

[2] Benedikt Ahrens, Krzysztof Kapulkin, and Michael Shulman. 2015. Univalent categories and the Rezk completion. *Mathematical Structures in Computer Science* 25, 05 (2015), 1010–1039. https://doi.org/10.1017/S0960129514000486

[3] Benedikt Ahrens, Paige Randall North, Michael Shulman, and Dimitris Tsementzis. 2020. A Higher Structure Identity Principle. (2020). arXiv:2004.06572

[4] Danil Annenkov, Paolo Capriotti, Nicolai Kraus, and Christian Sattler. 2019. Two-Level Type Theory and Applications. (2019). arXiv:1705.03307v3

[5] Georges Blanc. 1978. Équivalence naturelle et formules logiques en théorie des catégories. *Arch. Math. Logik Grundlag.* 19, 3-4 (1978), 131–137. https://doi.org/10.1007/BF02011874

[6] Thierry Coquand and Nils Anders Danielsson. 2013. Isomorphism is equality. *Indagationes Mathematicae* 24, 4 (2013), 1105 – 1120. https://doi.org/10.1016/j.indag.2013.09.002 In memory of N.G. (Dick) de Bruijn (1918–2012).

[7] Peter Freyd. 1976. Properties invariant within equivalence types of categories. In *Algebra, topology, and category theory (a collection of papers in honor of Samuel Eilenberg)*. Academic Press, New York, 55–61.

[8] Chris Kapulkin and Peter LeFanu Lumsdaine. 2019. The simplicial model of univalent foundations (after Voevodsky). *Journal of the European Mathematical Society* (2019). arXiv:1211.2851 To appear.

[9] Michael Makkai. 1995. First Order Logic with Dependent Sorts, with Applications to Category Theory. (1995). http://www.math.mcgill.ca/makkai/folds/foldsinpdf/FOLDS.pdf.

[10] Dimitris Tsementzis. 2017. A Higher Structure Identity Principle. (2017). arXiv:1702.07776

[11] The Univalent Foundations Program. 2013. *Homotopy Type Theory: Univalent Foundations of Mathematics.* https://homotopytypetheory.org/book, Institute for Advanced Study.

[12] Vladimir Voevodsky. 2015. An experimental library of formalized Mathematics based on the univalent foundations. *Mathematical Structures in Computer Science* 25, 5 (2015), 1278–1294. https://doi.org/10.1017/S0960129514000577