Categories as Semicategories with Identities

Joshua Chen, Tom de Jong, Nicolai Kraus, and Stiéphen Pradal

University of Nottingham, Nottingham, UK
{joshua.chen, tom.dejong, nicolai.kraus, stiephen.pradal}@nottingham.ac.uk

Motivation The development of category theory inside type theory has a long history, and many libraries for proof assistants such as Agda or Coq contain results on categories [4, 9, 10, 11, 12, 14]. In a type theory without UIP, in particular in HoTT, the theory of 1-categories is often not applicable for the study of types and more general (i.e., higher) notions of categories are required. For example, the universe of types and functions is adequately described as an $(\infty, 1)$-category. Unsurprisingly, already writing down the definition of such a higher category is involved and a careful approach to organising the huge number of components is needed.[1]

One approach to defining higher categories is to first consider the composition structure (i.e., morphisms, composition, associativity, Mac Lane's pentagon coherence, ...). This leads to a notion of higher semicategory. We then want to describe the higher categories as those higher semicategories that happen to have identities. If we can formulate "having identities" as a propositional property, the higher categories become a subtype of the higher semicategories.

In this talk, we present several different (equivalent) definitions of the property "having identities". Instead of higher categories, we work with a 1-categorical notion of semicategory, a "wild" (untruncated) and a priori ill-behaved concept that generalises both "honest" semicategories (with set-truncated morphism types) and $(\infty, 1)$-semicategories (with all coherences). The fact that this is possible is very fortunate as it simplifies the situation significantly compared to the ∞-categorical setting, but it of course leads to the question in which sense our identity structures are "correct" for $(\infty, 1)$-categories. We discuss this question at the end.

Notions of identities in wild semicategories A wild semicategory is a tuple $(\mathsf{Ob}, \mathsf{hom}, \circ, \alpha)$ where α witnesses associativity. The attribute *wild* indicates that we do not place a truncation condition on the family hom.

Naive identities A direct way to define an identity structure is to ask for a function $\mathsf{id} : \Pi_{x:\mathsf{Ob}} \mathsf{hom}(x, x)$ together with identity laws $\lambda_f : \mathsf{id}_y \circ f = f$ and $\rho_f : f \circ \mathsf{id}_x = f$. Since hom is not required to be a family of sets, this formulation of *having naive identities* is not a proposition and it does not automatically satisfy the coherences that one would expect of an identity in a higher category, such as $\lambda_{\mathsf{id}} = \rho_{\mathsf{id}}$. We write NaId_x for the type of triples $(\mathsf{id}_x, \lambda, \rho)$.

Idempotent equivalences A less direct but more well-behaved definition of an identity structure is to ask for an *idempotent equivalence* on each object ([7]; cf. the *weak units* of [5]). Here, a morphism f is an equivalence if both pre- and post-composition with f is an equivalence of types in the usual (HoTT) sense and we write $\mathsf{eqv}(x, y)$ for the subtype of $\mathsf{hom}(x, y)$ that are equivalences. A morphism $f : \mathsf{hom}(x, x)$ is idempotent if $f \circ f = f$. Clearly, we would expect an identity morphism to be both an equivalence and idempotent, and it turns out that this expectation can be reversed: an idempotent equivalence is always a naive identity in the above sense. This notion is well-behaved since the type $\mathsf{IdemEqv} :\equiv \Pi_{x:\mathsf{Ob}} \Sigma_{i:\mathsf{eqv}(x,x)}(i \circ i = i)$ is a proposition [7].

[1] It is a well-known open question, and one of the major unsolved problems of the field, whether homotopy type theory [13] is expressive enough to formulate the definition of an $(\infty, 1)$-category such that the universe is an instance. The difficulty is to find a way (or determine that there is no way) to encode the infinite number of morphism levels. 2LTT [1] is a setting in which this can be done. The current abstract is *not* on this issue.

Harpaz's identities Following an idea by Harpaz [3], we can ask that there is an equivalence out of each object x, that is, $\Sigma_{y:\mathsf{Ob}}\,\mathsf{eqv}(x,y)$. If we want this to be a proposition, we can truncate (i.e., replace Σ_y by $\exists_y$); this variation is still sufficiently strong to derive a naive identity structure. Alternatively, we can ask for the type of outgoing equivalences (the type of tuples (y,f,e)) to be contractible. It turns out that this version defines *univalent* identities [2]. We write $\mathsf{HarpazId} :\equiv \Pi_{x:\mathsf{Ob}}\exists_{y:\mathsf{Ob}}\,\mathsf{eqv}(x,y)$ and $\mathsf{uHarpazId} :\equiv \Pi_{x:\mathsf{Ob}}\,\mathsf{isContr}(\Sigma_{y:\mathsf{Ob}}\,\mathsf{eqv}(x,y))$.

Identities via (co)slices In category theory, the identity on x is the terminal (resp. initial) object in the slice over (resp. the coslice under) x. Reversing this, we get yet another method to characterise an identity structure in semicategories. Note that wild semicategories are not sufficiently well-behaved to construct slices or coslices as associativity cannot be derived (as explained in e.g. [7]). However, sufficient structure can be constructed to define what it means to be initial or terminal. After unfolding the definition, this leads to the simple definition $\mathsf{SliceId} :\equiv \Pi_{x:\mathsf{Ob}}\|\mathsf{eqv}(x,x)\|$.

Equivalence of the above notions For a semicategory with set-truncated families of morphisms, a naive identity structure is unique if it exists; in other words, NaId is a proposition. This is not the case for a wild semicategory, but we could explicitly truncate to get a proposition $\|\mathsf{NaId}\|$. By combining results from several papers we can then show:

Theorem 1. *For a given wild semicategory, the four types* $\Pi_{x:\mathsf{Ob}}\|\mathsf{NaId}_x\|$, $\mathsf{IdemEqv}$, $\mathsf{HarpazId}$, $\mathsf{SliceId}$ *are equivalent propositions.*

Proof. Three of the types are explicitly constructed to be propositions. In contrast, it is not automatic that $\mathsf{IdemEqv}$ is a proposition: While *being an equivalence* is a proposition, the *type* of equivalences is in general not a proposition, and neither is the statement that a morphism is idempotent. The result was shown by the third-named author in [7] and the strategy is to show that, if an identity-like morphism is given, then every idempotent equivalence has to be equal to it. We refer to the formalisation[2] for the details.

- $\Pi_{x:\mathsf{Ob}}\|\mathsf{NaId}_x\| \leftrightarrow \mathsf{IdemEqv}$ [7]: Naive identities are idempotent equivalences and vice versa.
- $\mathsf{HarpazId} \to \mathsf{IdemEqv}$: This uses an insight of Harpaz [3] in a type-theoretic setting. Given an equivalence $f : \mathsf{hom}(x,y)$, we can apply the inverse of $(f \circ \_)$ to f itself, and the result is an idempotent equivalence.
- Finally, $\mathsf{IdemEqv} \to \mathsf{SliceId} \to \mathsf{HarpazId}$ is easy. □

Discussion One approach to defining $(\infty,1)$-semicategories in a type-theoretic setting is to consider certain type-valued presheaves over the semi-simplex category Δ_+ [1, 2]. Morally, an identity structure corresponds to the maps present in the simplex category Δ but not in Δ_+. Unfortunately, the strategy of defining strict type-valued presheaves via type families only works for direct categories, which Δ is not. Approaches that include an identity structure include the use of a *direct replacement* of Δ [6, 8] or *homotopy coherent nerves* [8]; these structures consist of infinite towers of coherence data.

An $(\infty,1)$-semicategory $\mathcal{C}$ has an underlying wild semicategory $\mathcal{C}_1$. If $\mathcal{C}$ has an identity structure given by an infinite tower of coherence data, then $\mathcal{C}_1$ is trivially equipped with naive identities and thus any of the other discussed identity structures (apart from $\mathsf{uHarpazId}$, which is stronger). We conjecture that the converse holds as well; special cases of this expectation are verified in [2]. This conjecture would give us an easy way to construct the complete tower of coherences by checking any of the very easy conditions discussed above.

[2] **Formalisation** — browsable `html` version: `joshchen.io/agda/semicategories-with-identities/`; Agda source code: `github.com/jaycech3n/semicategories-with-identities`

References

[1] Danil Annenkov, Paolo Capriotti, Nicolai Kraus, and Christian Sattler. Two-level type theory and applications. arXiv[cs.LO]:1705.03307, 2019.

[2] Paolo Capriotti and Nicolai Kraus. Univalent higher categories via complete semi-Segal types. *Proceedings of the ACM on Programming Languages*, 2(POPL'18):44:1–44:29, 2017. Full version available at arXiv:1707.03693.

[3] Yonatan Harpaz. Quasi-unital ∞-categories. *Algebraic & Geometric Topology*, 15(4):2303–2381, 2015. doi:10.2140/agt.2015.15.2303.

[4] Jason Z. S. Hu and Jacques Carette. Agda-Categories: Category theory library for Agda. doi:10.1145/3410272, 2021.

[5] André Joyal and Joachim Kock. Coherence for weak units. *Documenta Mathematica*, 18:71–110, 2013.

[6] Joachim Kock. Weak identity arrows in higher categories. *International Mathematics Research Papers*, 2006:69163, 2006. doi:10.1155/IMRP/2006/69163.

[7] Nicolai Kraus. Internal ∞-categorical models of dependent type theory : Towards 2LTT eating HoTT. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–14, 2021. doi:10.1109/LICS52264.2021.9470667.

[8] Nicolai Kraus and Christian Sattler. Space-valued diagrams, type-theoretically (extended abstract). arXiv[math.LO]:1704.04543, 2017.

[9] Amélia Liao, Astra Kolomatskaia, and Reed Mullanix. 1lab. Available at https://1lab.dev/.

[10] The mathlib Community. The Lean mathematical library. *9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CCP'20)*, pages 367–381, 2020. doi:10.1145/3372885.3373824.

[11] Anders Mörtberg, Evan Cavallo, Felix Cherubini, Max Zeuner, Alex Ljungström, Andrea Vezzosi, et al. A standard library for Cubical Agda. Available at https://github.com/agda/cubical, 2018.

[12] Marco Perone et al. Idris category theory. Available at https://github.com/statebox/idris-ct.

[13] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. http://homotopytypetheory.org/book/, 2013.

[14] Vladimir Voevodsky, Benedikt Ahrens, Daniel Grayson, et al. UniMath — a computer-checked library of univalent mathematics. Available at http://unimath.org.