Final Coalgebra of the Finite Bag Functor

Philipp Joram and Niccolò Veltri

Department of Software Science, Tallinn University of Technology, Estonia

The powerset and multiset (or bag) functor, delivering the set of subsets (resp. multisubsets) of a given set, are fundamental mathematical tools in the behavioural analysis of nondeterministic systems. Such systems can be described as coalgebras $c : S \to F\,S$, with F being either powerset or bag functor, associating to each state $x : S$ the collection $c\,x$ of reachable states. When F is the bag functor, a state x can reach another state y in multiple ways, specified by the multiplicity of y in $c\,x$. In practice one often studies systems with finite reachable states and employs finite variants of the powerset and bag functor. The finite bag functor takes a set X to the set of finite bags of elements of X. In contrast to the finite powerset functor, these collections distinguish multiple occurrences of identical elements, e.g. $\{1, 2, 1\}$ is a bag containing an element 1 with multiplicity 2.

The behavior of a finitely nondeterministic system starting from a given initial state $x : S$ is fully captured by the final coalgebra of the collection functor F, whose elements are non-wellfounded trees obtained by iteratively running the coalgebra c on x. In recent work [8], we considered various constructions of the final coalgebra of the finite powerset functor in the setting of Homotopy Type Theory (HoTT). In this work, we explore possible definitions of the finite bag functor in HoTT and investigate whether they admit a final coalgebra.

In classical set theory, the final coalgebra of the finite bag functor F is obtained as the limit of the $\omega^{\mathbf{op}}$-chain [7, Ch. 10]

$$1 \xleftarrow{!} F1 \xleftarrow{F(!)} F^2 1 \xleftarrow{F^2(!)} F^3 1 \xleftarrow{F^3(!)} \cdots \tag{1}$$

where $F^n 1$ is the n-th iteration of the functor F on the singleton set 1, and $F^n(!)$ is the n-th iteration of the functorial action of F on the unique function ! targeting the set 1. In a constructive environment such as HoTT, it is not immediately obvious whether this construction still produces the final coalgebra or not.

We formalized our work in *Cubical Agda* [9], and the code is freely available at `https://github.com/phijor/agda-cubical-multiset`. Cubical Agda is a particular implementation of HoTT with support for the univalence principle ("equivalence of types is equivalent to equality of types", which is a provable theorem in Cubical Agda) and higher inductive types. A distinguishing feature of Cubical Agda is that the identity type on a type A is no longer inductively defined as in Martin-Löf Type Theory (MLTT). In particular, for proofs of identification $p, q : a =_A b$, the iterated identity type $p =_{a=_A b} q$ might be inhabited by more than one term, which in turn might have non-trivial identity type and so on. One says that A has *homotopy level n* if the n-th iterated identity type is trivial. We explicitly name the first instances of this hierarchy, and say that A is:

- $(n = 1)$ a *proposition*, if $\mathsf{isProp}\,A =_{\mathsf{df}} \forall(a, b : A).\ a = b$ is inhabited,
- $(n = 2)$ a *set*, if $\mathsf{isSet}\,A =_{\mathsf{df}} \forall(a, b : A).\ \mathsf{isProp}(a = b)$ is inhabited,
- $(n = 3)$ a *groupoid*, if $\mathsf{isGroupoid}\,A =_{\mathsf{df}} \forall(a, b : A).\ \mathsf{isSet}(a = b)$ is inhabited.

We stress that, when mentioning "sets" or "groupoids", we always refer to these definitions. In ordinary MLTT, the principle of uniqueness of identity proofs (UIP) implies that all types are sets. In HoTT, this is not true for arbitrary types.

Using *higher inductive types*, any type can be *truncated* to obtain a set:

Definition 1. *For any type A, the set-truncation $\|A\|_2$ is the type inductively defined by*

$$\frac{a : A}{|a|_2 : \|A\|_2} \qquad\qquad \frac{x, y : \|A\|_2 \qquad p, q : x = y}{\mathsf{squash}_2 : p = q}$$

This type is a "higher" inductive type since the rule squash_2 produces an identification $p = q$ instead of an element of $\|A\|_2$. By definition, $\mathsf{squash}_2 : \mathsf{isSet}\,\|A\|_2$. Similarly, the *propositional truncation* of A is the higher inductive type $\|A\|_1$ with a constructor $|\_|_1 : A \to \|A\|_1$ and a rule $\mathsf{squash}_1 : \mathsf{isProp}\,\|A\|_1$. Higher inductive types also allow us to model quotient types, in particular set-quotients:

Definition 2. *Given $A : \mathsf{Type}$ and a binary relation $(\sim) : A \to A \to \mathsf{Type}$, the* set-quotient *of A by $(\sim)$ is given by rules*

$$\frac{a : A}{[a]_2 : A /_2 \sim} \qquad \frac{a, b : A \qquad r : a \sim b}{\mathsf{eq}/_2 : [a]_2 = [b]_2} \qquad \frac{x, y : A /_2 \sim \qquad p, q : x = y}{\mathsf{squash}/_2 : p = q}$$

and forms a set by $\mathsf{squash}/_2$.

Finite Bags in Sets

In a first attempt, we define

$$\mathsf{FMSet}\,X =_{\mathsf{df}} \sum k : \mathbb{N}.\ (\mathsf{Fin}\,k \to X) /_2 \sim_k,$$

where $\mathsf{Fin}\,k$ is the type of numbers $< k$ and $(\sim_k)$ relates $v, w : \mathsf{Fin}\,k \to X$ iff there merely exists a permutation σ such that $v \circ \sigma = w$. So a finite bag consists of a number k (its size) and an equivalence class of k-elements of X, where the relation $(\sim_k)$ expresses that the order of the elements does not matter. $\mathsf{FMSet}\,X$ is a set, regardless of the homotopy level of X. Additionally, FMSet is invariant under set-truncation, i.e. $\mathsf{FMSet}\,\|X\|_2 \simeq \mathsf{FMSet}\,X$. The type of finite bags can equivalently be defined as the free commutative monoid on X, which can be directly expressed as a higher inductive type [3].

Trying to construct the final coalgebra of FMSet as the limit of the chain (1) (as traditionally done in a classical metatheory) is problematic. The first step in the construction would be showing that FMSet preserves ω-limits. In the case of the chain (1), this reduces to showing that the map $\mathsf{pres}\colon \mathsf{FMSet}(\lim_{n<\omega} \mathsf{FMSet}^n\,1) \to \lim_{n<\omega}(\mathsf{FMSet}^{n+1}\,1)$, defined via the universal property of the limit, is an equivalence of types. This is not the case in HoTT, since the latter is an inherently non-constructive principle:

Theorem 1. *The function $\mathsf{pres}\colon \mathsf{FMSet}(\lim_{n<\omega} \mathsf{FMSet}^n\,1) \to \lim_{n<\omega}(\mathsf{FMSet}^{n+1}\,1)$ is surjective, but its injectivity is equivalent to the* lesser limited principle of omniscience*, LLPO.*

LLPO [2, Ch. 1] is a weak version of the law of the excluded middle, and it is not provable from intuitionistic axioms alone. It states that, given an infinite stream of boolean values that yields true in at most one position, one can decide whether all even or all odd positions are false.

The non-constructive nature of the injectivity of pres can be attributed to the fact that the relation $(\sim_{sz})$ encodes permutations of multisets as property instead of data, and this makes it impossible to recover information about all terms in the chain.

We conclude from Theorem 1 that the classical construction of the final coalgebra for FMSet cannot be replicated in our constructive setting without assumption of classical principles. This is analogous to the case of the finite powerset functor, which we know to be suffering from similar issues [8].

Finite Bags in Groupoids

To remedy the situation, we introduce a bag functor that treats identifications of bags as data. Define the type of *(Bishop) finite sets* as

$$\mathsf{FinSet} =_{\mathsf{df}} \sum Y : \mathsf{Type}.\ \sum k : \mathbb{N}.\ \|Y \simeq \mathsf{Fin}\,k\|_1,$$

i.e. the type of all types merely equivalent to some $\mathsf{Fin}\,k$. While in this abstract we suppress any size-related issues (all the definitions and theorems given so far can be made universe-polymorphic), they play a crucial rôle in the formalization. Notice that FinSet is a large type compared to the types it ranges over, but it admits a small axiomatization as a higher inductive type Bij, originally introduced in [4]. Its introduction rules are the following (plus one stating that Bij is a groupoid):

$$\frac{n : \mathbb{N}}{\mathsf{obj}\,n : \mathsf{Bij}} \qquad \frac{m, n : \mathbb{N} \qquad \alpha : \mathsf{Fin}\,m \simeq \mathsf{Fin}\,n}{\mathsf{hom} : \mathsf{obj}\,m = \mathsf{obj}\,n} \qquad \frac{n : \mathbb{N}}{\mathsf{hom}(\mathrm{id}_{\mathsf{Fin}\,n}) = \mathsf{refl}}$$

$$\frac{m, n, o : \mathbb{N} \qquad \alpha : \mathsf{Fin}\,m \simeq \mathsf{Fin}\,n \qquad \beta : \mathsf{Fin}\,n \simeq \mathsf{Fin}\,o}{\mathsf{hom}(\beta \circ \alpha) = \mathsf{hom}\,\alpha \bullet \mathsf{hom}\,\beta}$$

Here, id is the identity-equivalence, $(\circ)$ composition of equivalences, $\mathsf{refl} : \mathsf{obj}\,n = \mathsf{obj}\,n$ a reflexivity proof of identity and $(\bullet)$ is transitivity of $(=)$.

With this in mind we define, for any type X, the type

$$\mathsf{Bag}\,X =_{\mathsf{df}} \sum x : \mathsf{Bij}.\ \langle x \rangle \to X,$$

where $\langle x \rangle : \mathsf{Type}$ is the type obtained from the equivalence $\mathsf{Bij} \simeq \mathsf{FinSet}$. Since Bij is a groupoid, each $\mathsf{Bag}\,X$ has a homotopy level of at least that of a groupoid. We conjecture that $\mathsf{Bag}\,X$ can be proved equivalent to the *free symmetric monoidal groupoid* on X defined as a HIT by Piceghello [6], which would serve as an alternative proof of MacLane's coherence for symmetric monoidal categories. In any case, we recover FMSet by truncating the higher structure of Bag:

Theorem 2. *For any type X, there is an equivalence $\|\mathsf{Bag}\,X\|_2 \simeq \mathsf{FMSet}\,X$.*

Following [5], we argue that this is the correct perspective on bags in HoTT, and substantiate the claim by the following unproblematic construction of the final coalgebra of Bag:

Theorem 3. *Let $L_{\mathsf{Bag}} =_{\mathsf{df}} \lim_{n<\omega} \mathsf{Bag}^n\,1$. The limit-preservation map* pres *is an equivalence of groupoids. In particular, the limit L_{Bag} is a fixpoint of* Bag *and its final coalgebra.*

This theorem is a direct consequence of a general result by Ahrens et al. [1], since Bag is a polynomial functor in groupoids and all polynomial functors admit a final coalgebra in HoTT, independently of the their homotopy level. When defined in terms of Bij, Bag is a small type family, and so is L_{Bag}. It would be a large type, had we defined it in terms of FinSet.

One might wonder whether the final coalgebra in groupoids can be used to define a final coalgebra also in sets. We are able to show that the set-truncation of the groupoid-limit $\|L_{\mathsf{Bag}}\|_2$ is a fixpoint of FMSet, i.e. $\mathsf{FMSet}\,\|L_{\mathsf{Bag}}\|_2 \simeq \|L_{\mathsf{Bag}}\|_2$. But further showing that this is the final coalgebra of the set-based bag functor FMSet again appears to be non-constructive:

Theorem 4. *Assuming the (full) axiom of choice, $\|L_{\mathsf{Bag}}\|_2$ is the final coalgebra of* FMSet.

Acknowledgments

This work was supported by the Estonian Research Council grant PSG749.

References

[1] Ahrens, B., Capriotti, P., Spadotti, R.: Non-wellfounded trees in Homotopy Type Theory. In: Proceedings of TLCA 2015. LIPIcs, vol. 38, pp. 17–30 (Apr 2015). doi:10.4230/LIPICS.TLCA.2015.17

[2] Bridges, D., Richman, F.: Varieties of Constructive Mathematics. Cambridge University Press (1987-04). doi:10.1017/cbo9780511565663

[3] Choudhury, V., Fiore, M.: Free commutative monoids in homotopy type theory (Oct 2021). doi:10.48550/arXiv.2110.05412

[4] Finster, E., Mimram, S., Lucas, M., Seiller, T.: A cartesian bicategory of polynomial functors in homotopy type theory. EPTCS 351, 2021, pp. 67-83 (Dec 2021). doi:10.4204/EPTCS.351.5, `https://github.com/smimram/fibred-polynomials`

[5] Kock, J.: Data Types with Symmetries and Polynomial Functors over Groupoids. Electronic Notes in Theoretical Computer Science **286**, 351–365 (Sep 2012). doi:10.1016/j.entcs.2013.01.001

[6] Piceghello, S.: Coherence for Monoidal and Symmetric Monoidal Groupoids in Homotopy Type Theory. Ph.D. thesis, University of Bergen, Norway (2021), `https://hdl.handle.net/11250/2830640`

[7] Rutten, J.J.M.M.: Universal coalgebra: a theory of systems. Theoretical Computer Science **249**(1), 3–80 (2000). doi:10.1016/S0304-3975(00)00056-6

[8] Veltri, N.: Type-theoretic constructions of the final coalgebra of the finite powerset functor. In: 6th International Conference on Formal Structures for Computation and Deduction, FSCD 2021, July 17-24, 2021, Buenos Aires, Argentina (Virtual Conference). Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2021). doi:10.4230/LIPICS.FSCD.2021.22

[9] Vezzosi, A., Mörtberg, A., Abel, A.: Cubical agda: a dependently typed programming language with univalence and higher inductive types. Proceedings of the ACM on Programming Languages **3**(ICFP), 1–29 (2019-07). doi:10.1145/3341691