

Department of Computer
Science and Technology

May 22, 2025

Efficient Coinductives through State-Machine Corecursors

William Sørensen

Gonville & Caius College

Submitted in partial fulfilment of the requirements for the
Computer Science Tripos, Part III

Declaration

I, William Sørensen of Gonville & Caius College, being a candidate for the course, hereby declare that this report and the work described in it are my own work, unaided except as may be specified below, and that the report does not contain material that has already been used to any substantial extent for a comparable purpose.

Signed: William Sørensen

Date: May 22, 2025

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliquam quaerat voluptatem. Ut enim aequo doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit.

Acknowledgements

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliquam quaerat voluptatem. Ut enim aequo doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distingue possit, augeri amplificarique non possit.

Contents

1	Introduction	6
1.1	Dependent type theory	6
1.2	Polynomial functors	6
1.2.1	Lean formalization	6
1.2.2	F-(co)algebras	6
2	Preparation	7
3	Implementation	8
3.1	The ABI Type	8
3.1.1	Weak univalence	9
3.1.2	Relation to Shrink and further universe transforming types	9
3.2	Stream implementation	9
3.3	Expanding the progressive approximation theory	9
3.3.1	Universe lifting of polynomial functors	10
3.3.2	Generalizing the corecursor	10
3.4	State machine encoding	10
3.4.1	PreM	10
3.5	Proving the equivalence	11
3.6	Cofix implementation	11
4	Evaluation	12
4.1	Performance between SME and PA	12
5	Conclusions	13
6	Appendices	14

Chapter 1

Introduction

1.1 Dependent type theory

1.2 Polynomial functors

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliquam quaerat voluptatem. Ut enim aequo doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos iridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et.

1.2.1 Lean formalization

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliquam quaerat voluptatem. Ut enim aequo doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos iridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et.

1.2.2 F-(co)algebras

Chapter 2

Preparation

Chapter 3

Implementation

3.1 The ABI Type

The problem the ABI type tries to tackle is one of abstracting the runtime datatype through functions. Given an isomorphism $\text{eq} : \alpha \cong \beta$ for some types α and β , my first try at solving this involved constructing an object $\text{ABI } \alpha\beta$, making the following commute:

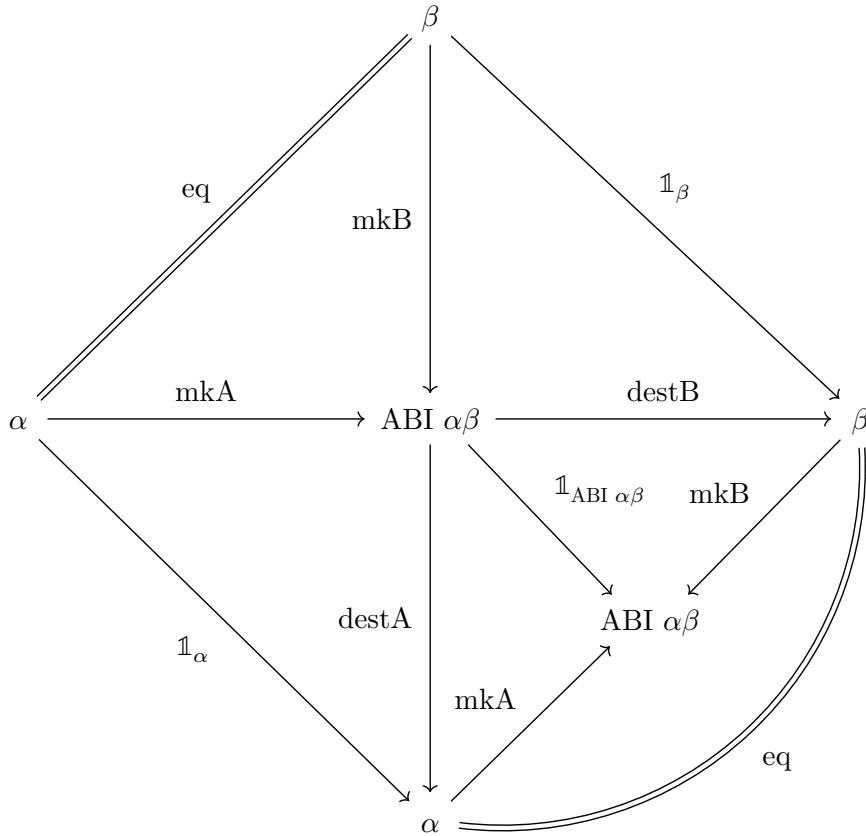


Figure 1: Operations on ABI

Additionally I had an elimination principle satisfying the two equations below.

```
1  def elim : {motive : ABI _ _ eq → Type w}
2      → (hLog : (z : A) → motive (mkA z))
```

Lean

```

3      → (hCheap : (z : B) → motive (mkB z))
4      → (eqA : ∀ z, hLog z ≈ hCheap (eq z))
5      → (eqB : ∀ z, hCheap z ≈ hLog (eq.symm z))
6      → (v : ABI _ _ eq) → motive v := _
7
8 theorem elimLog : {motive : carry → Type w}
9      → {hLog : (z : A) → motive (mkA z)}
10     → {hCheap : (z : B) → motive (mkB z)}
11     → {eqA : ∀ z, hLog z ≈ hCheap (eq z)}
12     → {eqB : ∀ z, hCheap z ≈ hLog (eq.symm z)}
13     → ∀ z, elim hLog hCheap eqA eqB (mkA z) = (hLog z) := _
14 theorem elimCheap : {motive : carry → Type w}
15     → {hLog : (z : A) → motive (mkA z)}
16     → {hCheap : (z : B) → motive (mkB z)}
17     → {eqA : ∀ z, hLog z ≈ hCheap (eq z)}
18     → {eqB : ∀ z, hCheap z ≈ hLog (eq.symm z)}
19     → ∀ z : B, elim hLog hCheap eqA eqB (mkB z) = (hCheap z) := _

```

Through quite a bit of work (which I call transliteration, as seen in the ABI file), You can free a universe level and open for a more general usage of the function.

3.1.1 Weak univalence

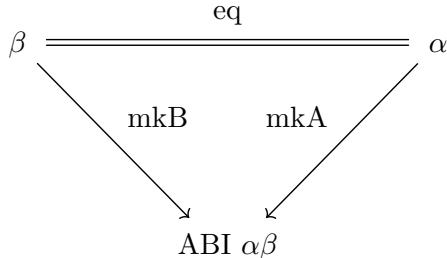


Figure 2: Weak univalence up to shrink

3.1.2 Relation to Shrink and further universe transforming types

3.2 Stream implementation

3.3 Expanding the progressive approximation theory

During the pheasability assesment I noticed that, in the current formalised theory of polynomials, the statement wouldn't even type-check. This stemmed from a problem with the corecursive principle for the M type in the old implementation. $\text{corec} : \{\alpha : \text{TypeVec}.\{\mathcal{U}\}n\} \rightarrow \{\beta : \text{Type } \mathcal{U}\} \rightarrow (g : \beta \rightarrow P(\alpha :: \beta)) \rightarrow \beta \rightarrow M\alpha^1$. The problem here is that both α and β have to both reside in \mathcal{U} . Solving this is done through the next two sections.

¹<https://github.com/leanprover-community/mathlib4/blob/7a60b315c7441b56020c4948c4be7b54c222247b/Mathlib/Data/PFunctor/Multivariate/M.lean#L152-L154>

3.3.1 Universe lifting of polynomial functors.

The main problem caused here comes from the fact that lean isn't cumulative. This means it is impossible to express a universe heterogeneous typevector. In other words $\alpha :: \beta$ is only typable if $\alpha : \text{TypeVec}.\{\mathcal{U}\}n$ and $\beta : \text{Type } \mathcal{U}$. The natural way of solving this is using the supremum in universe levels you get from $\text{ULift} : \text{Type } \mathcal{U} \rightarrow \text{Type } (\max \mathcal{U} \mathcal{V})$. This means we can have $\beta : \text{Type } \mathcal{U}$ and $\alpha : \text{Type } \mathcal{V}$, then lift both of them to a common universe $\text{ULift } \alpha :: \text{ULift } \beta : \text{TypeVec}.\{\max \mathcal{U} \mathcal{V}\}(n + 1)$ ².

Noticeable the next hurdle we encounter is that Pfunctors are restricted to a universe level. Recall the definition from Section 1.2.1. Observe how for a $\text{MvPfunctor}.\{\mathcal{U}\}n$, we require that both the head and child reside in \mathcal{U} . This will also cause problems, as looking back at the definition of the corecursor, we will require P to be able to accept $\text{ULift } \alpha :: \text{ULift } \beta$. If we do not add the ability to lift P , the unifier will force $\mathcal{U} = \mathcal{V}$, thereby invalidating all the work we did in the previous section. Luckily lifting a Pfunctor is relatively easy. We define it as $\text{ULift } P \triangleq \langle \text{ULift } P.1, \lambda x \mapsto \text{ULift } (P.2x) \rangle$. This works and now we can move on to our goal³.

3.3.2 Generalizing the corecursor

Now with all the work in the previous section, by generalizing corec⁴, we can define $\text{corecU} : \{\alpha : \text{TypeVec}.\{\mathcal{U}\}n\} \rightarrow \{\beta : \text{Type } \mathcal{V}\} \rightarrow (g : \beta \rightarrow \text{ULift } P(\text{ULift } \alpha :: \text{ULift } \beta)) \rightarrow \beta \rightarrow M.\{\mathcal{U}\}\alpha$. Notably we are able to fit the object into \mathcal{U} (as opposed to in the SME).

The expected diagram using corecU and dest commutes.

3.4 State machine encoding

Noting the definition of corecU, one might wonder if you could define M from first principles for this. The problem one encounters is one of universes. As seen in the definition above, if one were to define a type whose constructor is directly the corecU definition, it would hold a $\beta : \text{Type } \mathcal{V}$. This then forces the object to reside in $\text{Type } \max \mathcal{U}(\mathcal{V} + 1)$. This is a problem as one loses most closure results as you will be lifting more and more. The main benefit from this is the performance aspect. Going from reading to a depth n , to a depth $n + 1$ is not $\mathcal{O}(1)$ instead of $\mathcal{O}(n)$. This will be seen in Section 4.1. We will henceforth refer to the datatype SME.PreM.

3.4.1 PreM

As we speak about in Section 1.2.2, the M Type is the terminal object in the category of coalgebras. We can see through reasoning (cumulatively) in this category that PreM is weakly terminal. Looking at this category we want to somehow force the incoming morphisms together. This corresponds exactly to quotienting, for this we will use Bisimilarity.

²Note we overload ULift as a notation to refer to lifting TypeVecs as well

³TODO: Speak with JV / W to see if this might be done in the lit, Ex : Locally presentable and accessible categories Adamek Rosicki

⁴Done in PR NUMBER

3.5 Proving the equivalence

3.6 Cofix implementation

Chapter 4

Evaluation

4.1 Performance between SME and PA

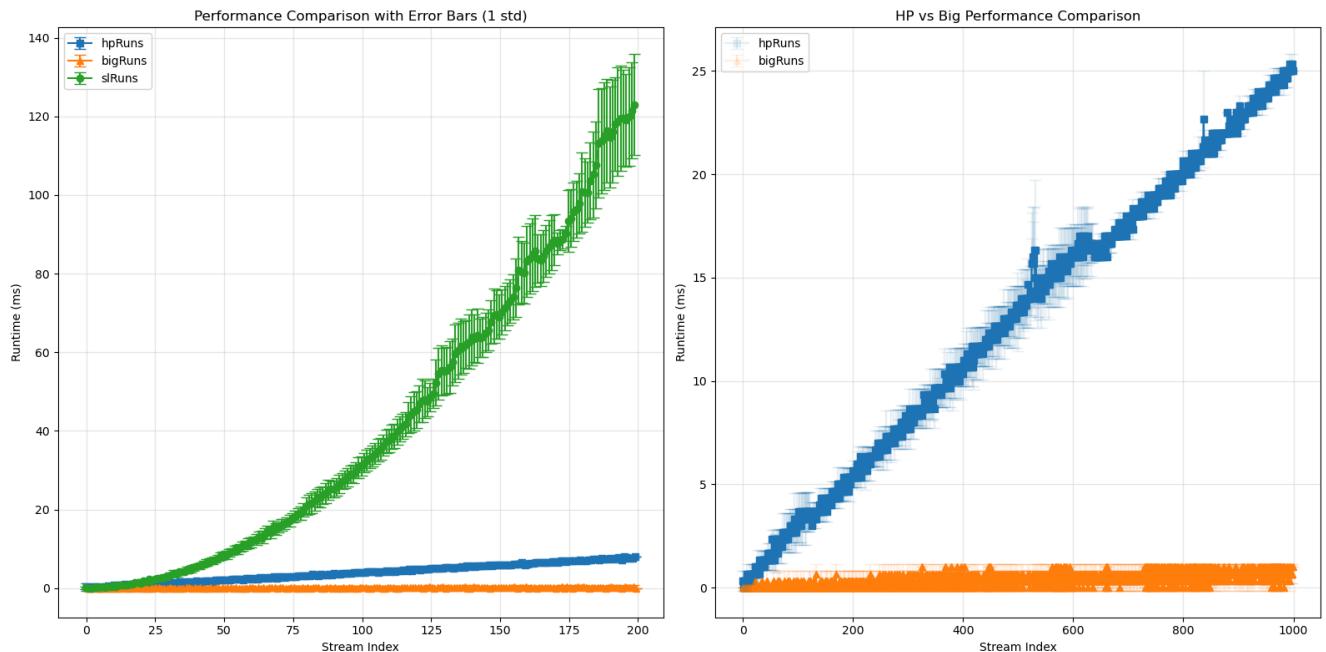


Figure 3: Performance of SME, PA, and Big representation

Chapter 5

Conclusions

Chapter 6

Appendices