A Categorical Representation of Thin Trees

Anton Chernev[1], Corina Cîrstea[2], Alexandre Goy[2], Helle Hvid Hansen[1], and Clemens Kupke[3]

[1] University of Groningen, Netherlands
[2] University of Southampton, United Kingdom
[3] University of Strathclyde, United Kingdom

Abstract

Infinite trees with countably many branches, called thin trees, have been studied via methods from automata theory and algebra. We take a categorical approach to thin trees using the framework of coalgebra. We show that the collection of thin trees can be seen as an initial algebra satisfying a certain axiom. We prove this by defining an algebra of thin tree representatives and showing that each thin tree has a canonical representative.

1 Background

Infinite words serve as a basis for the analysis of reactive systems and have been studied by means of automata and algebraic recognition [6]. Subsequently, infinite trees have also become an object of study, as they can express more complex systems where branching occurs. Tools from automata theory have been adapted to infinite trees [4].

In [7, 5] the authors look into automata and algebras for a class of infinite trees called *thin trees*. These are the trees that have countably many infinite branches. Every node in a thin tree can be assigned an ordinal called *rank*, which allows for inductive reasoning on the rank of thin trees. Moreover, languages of thin trees admit an algebraic characterisation via *thin algebras*, which are an extension of the notion of an ω-semigroup for infinite words. Thin algebras and induction on the rank are used to prove that languages of thin trees are recognised by *unambiguous automata*, i.e., automata that have unique accepting runs.

2 Contribution

In our current work we employ category theory to provide a uniform account of thin trees for a finite signature F. We base our approach on the formalisation of trees via F*-coalgebras* for a polynomial functor F over Set (see, e.g., [3]). Indeed, every tree with branching type F corresponds to an element of the final F-coalgebra (Z, ζ), and every element of (Z, ζ) can be unravelled into a tree. We take a look into the subcoalgebra (Z^{Th}, ζ^{Th}) of (Z, ζ) consisting of those elements whose unravelling is a thin tree. By endowing Z^{Th} with a suitable algebraic structure β^{Th}, we characterise (Z^{Th}, β^{Th}) as the initial object in a certain category of algebras ThAlg. In this way, we capture the principle of induction on the rank of thin trees via the universal property of initiality. Moreover, objects in ThAlg allow for algebraic recognition of languages of thin trees, analogously to thin algebras in [7]. This paves the way for future work on categorifying properties of thin trees, such as the existence of unambiguous automata.

3 Universal Property of Thin Trees

Here we give some details behind the construction of (Z^{Th}, β^{Th}). A key ingredient is the *functor derivative* F' [1], which represents the type of *contexts*, i.e., tree nodes where one successor is replaced by a hole. A context c can be composed with a tree or another context c' by plugging c' into the hole of c.

We define the type of streams of F-contexts $G \coloneqq (F')^\omega$ and denote the initial $(F+G)$-algebra by (A, α). Every term in A can be seen as a representative of a tree. If a term has type F, we interpret it as a tree node with given immediate successors. If it has type G, we interpret it as the tree obtained by composing infinitely many F-contexts. This gives rise to an *interpretation map* $int : A \to Z$. Moreover, we observe that interpretations are thin, i.e., $int[A] \subseteq Z^{Th}$. However, one element of Z^{Th} can have many representatives in A. For example, consider an element of Z^{Th} whose unravelling consists of a single infinite branch. It can be represented as a stream x_1 of contexts whose only successor is the hole, or as a node x_2 whose only successor is the stream x_1.

In order to get unique representatives, we quotient (A, α) by the congruence $\approx$ generated by the following axiom (†). We identify a term x of type G with the term y of type F obtained by plugging $tail(x)$ into the context $head(x)$. For instance, (†) will directly identify x_1 and x_2 from the example above. We show that $\approx$ is sound for int, i.e., terms identified by $\approx$ have the same interpretation. In order to show that each element of Z^{Th} is represented by a unique equivalence class of $\approx$, we introduce the notion of a *normal term*. It is defined via the *rank* of a term, which is the earliest step in the initial colimit construction of (A, α) at which the term appears. Now a term is called normal if it has the least rank among the terms with the same interpretation, and all its subterms are normal.

We prove two main results: (1) each element of Z^{Th} has a unique normal representative, and (2) the quotient relation $\approx$ identifies each term with its corresponding normal term. As a result, we conclude that the quotient of (A, α) by $\approx$ contains a unique representative for each element of Z^{Th}. Thus, for a suitable $(F+G)$-algebra structure β^{Th}, we have that (Z^{Th}, β^{Th}) is isomorphic to the quotient of (A, α) by $\approx$, so (Z^{Th}, β^{Th}) is initial among all $(F+G)$-algebras satisfying the axiom (†).

Beyond the purposes of our proofs, we hope that thin tree representatives can find use in applications, such as automata learning, where algorithms are sensitive to the particular presentation of objects [2].

References

[1] M. Abbott, T. Altenkirch, C. McBride, and N. Ghani. ∂ for data: Differentiating data structures. *Fundamenta Informaticae*, 65:1–28, 01 2005.

[2] Dana Angluin and Dana Fisman. Learning regular omega languages. *Theoretical Computer Science*, 650:57–72, 2016. Algorithmic Learning Theory.

[3] Corina Cîrstea and Clemens Kupke. Measure-theoretic semantics for quantitative parity automata. In Bartek Klin and Elaine Pimentel, editors, *31st EACSL Annual Conference on Computer Science Logic*, Leibniz International Proceedings in Informatics, pages 14:1–14:20, February 2023.

[4] E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*. Springer-Verlag, Berlin, Heidelberg, 2002.

[5] T. Idziaszek, M. Skrzypczak, and M. Bojańczyk. Regular languages of thin trees. *Theory of Computing Systems*, 20, 04 2015.

[6] D. Perrin and J.-E. Pin. *Infinite Words: Automata, Semigroups, Logic and Games*, volume 141 of *Pure and applied mathematics*. Elsevier, 2004.

[7] M. Skrzypczak. *Descriptive Set Theoretic Methods in Automata Theory - Decidability and Topological Complexity*, volume 9802 of *Lecture Notes in Computer Science*. Springer, 2016.