Unambiguous Acceptance of Thin Coalgebras

Anton Chernev[a,1]   Corina Cîrstea[b,2]   Helle Hvid Hansen[a,3]   Clemens Kupke [c,4]

[a] *University of Groningen*
Groningen, Netherlands

[b] *University of Southampton*
Southampton, United Kingdom

[c] *University of Strathclyde*
Glasgow, United Kingdom

Abstract

Automata admitting at most one accepting run per structure, known as unambiguous automata, find applications in verification of reactive systems as they extend the class of deterministic automata whilst maintaining some of their desirable properties. In this paper, we generalise a classical construction of unambiguous automata from thin trees to thin coalgebras for analytic functors. This achieves two goals: extending the existing construction to a larger class of structures, and providing conceptual clarity and parametricity to the construction by formalising it in the coalgebraic framework. As part of the construction, we link automaton acceptance of languages of thin coalgebras to language recognition via so-called coherent algebras, which were previously introduced for studying thin coalgebras. This link also allows us to establish an automata-theoretic characterisation of languages recognised by finite coherent algebras.

Keywords: Coalgebra, unambiguous automaton, thin tree, thin coalgebra, verification.

1 Introduction

Background and Motivation Model checking [3] of reactive systems makes extensive use of automata over infinite objects [8]. A core result facilitating the use of infinite word automata in verification is the determinisation of parity automata. However, this result is limited to infinite words, so systems whose runs exhibit tree-like structure call for more refined theoretical tools.

Recent work [7,6] (building on [17]) presents coalgebraic approaches to quantitative model checking using parity automata. Coalgebra [15] allows for a unified treatment of various system types by viewing these as coalgebras for a functor describing the system type. In particular, [6] proposes an approach to quantitative model checking of systems with quantitative branching type given by a monad T and the structure of system runs given by a polynomial functor F. A key condition in [6] is that the property to be checked must be given by an F-coalgebra automaton [12] that is *unambiguous*, i.e., there is at most one accepting run on each coalgebra. This raises the question of when an equivalent unambiguous F-coalgebra

automaton can be constructed from a nondeterministic one. This question is also of fundamental interest and the coalgebraic framework allows to investigate for which system types unambiguous acceptance results can be obtained. For ordered ranked trees, there are languages that are not accepted by an unambiguous automaton [4]. However, for the subclass of *thin trees*, i.e., trees with only countably many infinite branches, [16] shows how to construct from a nondeterministic automaton, an automaton that unambiguously accepts the same thin trees. The construction goes via *thin algebras*: every automaton can be transformed into a finite thin algebra, which can be transformed into an automaton that is unambiguous on thin trees.

Inspired by these results on thin trees, we showed in [5] that thin trees and their inductive characterisation can be generalised to the level of F-coalgebras for an analytic functor F. Analytic functors [11] include polynomial functors (the type of ordered ranked trees) and quotients thereof such as the bag functor. In the present paper, we build on the algebraic characterisation from [5] of thin F-coalgebras via so-called *coherent algebras* in order to prove unambiguous acceptance for thin F-coalgebras for analytic F.

Contributions We summarise our contributions below.

- We show that, when restricting to thin F-coalgebras for analytic functors F, every (nondeterministic) F-coalgebra automaton can be transformed into an equivalent unambiguous F-coalgebra automaton. We thus extend the results for thin trees [16], thereby making a step towards applications in quantitative model checking [6].

- We give an automata-theoretic characterisation of languages recognised by finite coherent algebras; these are precisely the languages accepted by F-automata with a so-called *prefix-agnostic* acceptance condition, which informally means that acceptance does not depend on any finite prefix of paths in the run.

- When instantiated to a polynomial functor F, our unambiguous automaton construction provides a categorical account of the classical construction in [16]. In particular, thin algebras arise as coherent algebras with additional structure, which we call *rational coherent algebras*.

We obtain these results as follows. In order to define unambiguous acceptance, in Section 3, we generalise the concept of run for F-coalgebra automata in [6] from polynomial F to analytic F. In Section 4, we show how to transform an automaton into a finite coherent algebra recognising the same language restricted to thin coalgebras. This construction works not just for parity automata, but, more generally, for automata with a prefix-agnostic acceptance condition. We identify *rational coherent algebras* as the coherent algebras obtained from parity automata. In Section 5, we show how to transform a finite coherent algebra into an automaton, called the *algebraic automaton*, which unambiguously accepts precisely the thin coalgebras that are recognised by the algebra. In order to prove correctness of this construction, we show that runs of the algebraic automaton correspond to certain coalgebra-to-algebra morphisms called *markings*. The uniqueness of markings, and hence of runs, follows from thin coalgebras being recursive thanks to their inductive structure. Finally, in Section 6, we combine the two constructions to obtain our main result, the transformation of an automaton into an automaton which, over thin coalgebras, is unambiguous and equivalent to the original one. In addition, we show that the languages recognised by finite coherent algebras coincide with the languages of thin behaviours accepted by automata with prefix-agnostic acceptance.

We finish the section with a brief example of the significance of our unambiguous automaton construction for model checking. Figure 1 depicts (a variant of) the simple probabilistic server from [6]. The state diagram on the left consists of a server and a worker. At each step, the server process spawns a worker with probability $\frac{1}{5}$ and returns to itself. A worker process performs a computation with probability $\frac{7}{8}$ and finishes otherwise. The type of this system is given by the functor $T \circ F$ where T is the distribution monad and F is a polynomial functor with a binary operation $fork$, two unary operations $wait, compute$ and a nullary operation $done$. On the right we see a possible execution (or trace) of the system. Suppose we are given a property P of system executions, such as "there exists a worker that never finishes". The framework [6] can then determine the probability with which P holds, as long as P is specified by an automaton that has at most one accepting run on each possible execution. Consider the automaton $\mathcal{A}$ for P that guesses at each $fork$ whether the worker does not terminate. This automaton is ambiguous, so it cannot be readily used for determining the probability. Note, however, that all executions of the given system are thin, because the server can spawn at most countably many workers. Therefore we can apply our unambiguous automaton construction to $\mathcal{A}$, so that the resulting automaton satisfies the desired condition of having at most one accepting run for each possible execution.
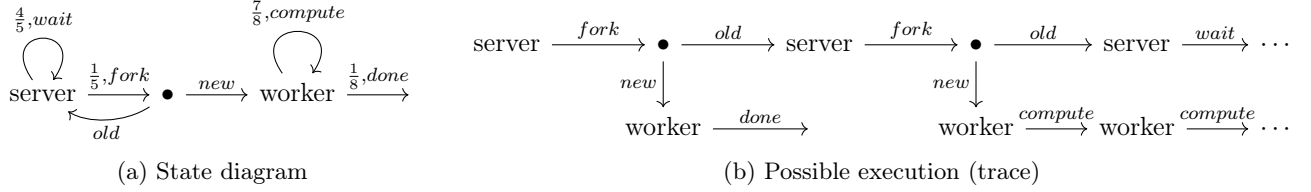
(a) State diagram

(b) Possible execution (trace)

Fig. 1. Example of a probabilistic server

2 Preliminaries

2.1 Automata and Algebras for Languages of Infinite Words

We begin by reviewing basics from the classical theory of automata on infinite words [8]. There are multiple types of equivalent infinite word automata, but here we focus on *(nondeterministic) parity word automata*. Given a finite *alphabet* Σ, a nondeterministic parity word automaton is a tuple $\mathcal{A} = (Q, \delta, Q_I, \Omega)$, where Q is a finite set of states, $\delta : Q \to \mathcal{P}(\Sigma \times Q)$ is a transition function, $Q_I \subseteq Q$ is a set of initial states and $\Omega : Q \to \omega$ is a priority function. An *accepting run* of $\mathcal{A}$ on an infinite word $x = (a_n)_{n\in\omega} \in \Sigma^\omega$ is a sequence of states $(q_n)_{n\in\omega} \in Q^\omega$ such that $q_0 \in Q_I$, $(a_n, q_{n+1}) \in \delta(q_n)$ and $\limsup_{n\in\omega} \Omega(q_n)$ is even, i.e., the largest priority occurring infinitely often is even. An infinite word x is *accepted* by $\mathcal{A}$ if there exists an accepting run of $\mathcal{A}$ on x. Languages (i.e., sets) of infinite words accepted by a nondeterministic parity word automaton are called *ω-regular*. An automaton $\mathcal{A}$ is *deterministic* if Q_I is a singleton and for each $q \in Q$ and $a \in \Sigma$, we have a single pair $(a, q_1) \in \delta(q)$. For convenience, we write deterministic parity word automata as (Q, δ, q_I, Ω) where $\delta : Q \to Q^\Sigma$ and $q_I \in Q$. An important result is that deterministic parity word automata accept the same languages as all (nondeterministic) parity word automata.

An alternative, algebraic approach to characterising ω-regular languages is via *ω-semigroups* [14, Chapter 2]. An ω-semigroup is a two-sorted algebraic structure (V, W) with three operations $\cdot : V \times V \to V$, $\times : V \times W \to W$, $\Pi : V^\omega \to W$, satisfying certain associativity axioms. In order to get some intuition about ω-semigroups, consider $(\Sigma^+, \Sigma^\omega)$, which is the ω-semigroup freely generated by Σ. Here $\cdot$ is concatenation between two finite words, $\times$ is concatenation between a finite and an infinite word and Π is concatenation of infinitely many finite words. A *homomorphism* between ω-semigroups (V_1, W_1) and (V_2, W_2) is a pair of maps $f = (f_V, f_W)$, where $f_V : V_1 \to V_2$, $f_W : W_1 \to W_2$, that preserves the ω-semigroup operations. The key property of ω-semigroups is that $L \subseteq \Sigma^\omega$ is ω-regular if and only if there exists a finite ω-semigroup (V, W), a homomorphism $f : (\Sigma^+, \Sigma^\omega) \to (V, W)$ and a *recognising set* $U \subseteq W$ such that $L = f_W^{-1}(U)$.

There exist extensions of parity automata from words to other infinite structures, such as binary trees. Instead of considering automata running on some concrete structures, we will work with *F-coalgebra automata* (see Section 3) that run on *F-coalgebras*.

2.2 F-Coalgebras and F-Algebras

F-coalgebras [15] are a formalism for modelling state-based systems that is parametric in the transition type F. Let F be an endofunctor on the category Set. An *F*-coalgebra is a tuple (X, ξ) consisting of an object X and a morphism $\xi : X \to FX$. An *F-coalgebra morphism* $f : (X, \xi) \to (Y, \upsilon)$ is a map $f : X \to Y$ (in Set) such that $\upsilon \circ f = Ff \circ \xi$. Informally, *F*-coalgebra morphisms map states in such a way that the transition structure is preserved. *F*-coalgebras, together with *F*-coalgebra morphisms, form a category. A terminal object (Z, ζ) in this category is called a *final F-coalgebra* and its elements can be thought of as *abstract behaviours*. By selecting a *root state* x_I in a coalgebra (X, ξ), we get a *pointed F-coalgebra* (X, ξ, x_I). *Pointed F-coalgebra morphisms* are *F*-coalgebra morphisms that also preserve the root.

Given F-coalgebras (X, ξ) and (Y, υ), two states $x \in X$ and $y \in Y$ are *behaviourally equivalent* if there exist *F*-coalgebra morphisms $f_X : (X, \xi) \to (W, \eta)$ and $f_Y : (Y, \upsilon) \to (W, \eta)$ into a third *F*-coalgebra such that $f_X(x) = f_Y(y)$. Two pointed *F*-coalgebras (X, ξ, x_I), and (Y, υ, y_I) are behaviourally equivalent if x_I and y_I are behaviourally equivalent. Under the assumption that F preserves weak pullbacks, behavioural equivalence amounts to the existence of a *span* of pointed coalgebra morphisms, i.e., a pointed *F*-coalgebra (R, ρ, s_I) with pointed morphisms $f_X : (R, \rho, s_I) \to (X, \xi, x_I)$ and $f_Y : (R, \rho, s_I) \to (Y, \upsilon, y_I)$.

Assuming that F preserves intersections and preimages, there exists a natural transformation $\text{Base}_F :$

3

$F \Rightarrow \mathcal{P}$, where $\mathcal{P}$ is the covariant power-set functor (see [9, Theorem 8.1]). For $\bar{x} \in FX$, $\mathrm{Base}_F(\bar{x}) \subseteq X$ is the least set such that $\bar{x} \in F(\mathrm{Base}_F(\bar{x}))$. The notion of base allows us to define *reachable* pointed F-coalgebras. These are pointed coalgebras (X, ξ, x_I) where for every $x \in X$, there exists a finite sequence $x_0, x_1, \ldots, x_n$ such that $x_0 = x_I$, $x_n = x$ and $x_{i+1} \in \mathrm{Base}_F(\xi(x_i))$ for all $i < n$. As the name suggests, every state in a reachable coalgebra can be reached from the root along some transitions. One readily observes that reachable coalgebras come with an induction principle: if $P \subseteq X$ is a property such that $x_I \in P$ and, for all $x \in X$, $x \in P$ implies $\mathrm{Base}(\xi(x)) \subseteq P$, then $P = X$.

F-algebra is the dual notion of F-coalgebra. An F-algebra is a pair (C, γ) with $\gamma : FC \to C$. An *F-algebra morphism* $f : (B, \beta) \to (C, \gamma)$ is then a map $f : B \to C$ with $f \circ \beta = \gamma \circ Ff$. An *initial F-algebra* is an initial object in the category of F-algebras and F-algebra morphisms. An F-algebra can be thought of as an algebra with a (generalised) signature F, and the elements of an initial F-algebra can be seen as equivalence classes of terms.

Given an F-coalgebra (X, ξ) and an F-algebra (C, γ), an *F-coalgebra-to-algebra morphism* is a map $f : X \to C$ satisfying $f = \gamma \circ Ff \circ \xi$. An F-coalgebra (X, ξ) is *recursive* if for every F-algebra (C, γ), there exists a unique F-coalgebra-to-algebra morphisms from (X, ξ) to (C, γ). Recursive coalgebras capture the idea of recursion on well-founded relations (see [2] for details).

In this paper, we will work with coalgebras for *analytic functors*.

2.3 Analytic Functors

Analytic functors [11] (see also [10]) generalise polynomial functors by allowing symmetries of successors, thus including, for instance, the bag functor. They were shown in [5] to be a natural setting for studying *thin coalgebras* (see Section 2.4). While here we give the basic definitions, we refer the reader to [5, Sections II,III] for a more detailed discussion with examples.

Given sets X, U and a group H of permutations on U, H *acts* on the set X^U of functions by $\sigma \cdot \phi = \phi \circ \sigma^{-1}$, for $\sigma \in H$ and $\phi \in X^U$. The set of *orbits* of this action is written as X^U/H, with elements of the form $[\phi]_H = \{\psi \in X^U \mid \exists \sigma \in H(\psi = \sigma \cdot \phi)\}$. An analytic functor is a functor of the form $F(X) = \bigsqcup_{i \in I} X^{U_i}/H_i$ where I is an index set, U_i is a finite set and H_i is a group of permutations on U_i, for all $i \in I$. Thus elements of $F(X)$ are of the form $(i, [\phi]_{H_i})$. We think of the sets U_i as positions to which we assign data in X. These positions can be permuted according to H_i. For a function $f : X \to Y$, $F(f)(i, [\phi]_{H_i}) = (i, [f \circ \phi]_{H_i})$.

We will use the notion of *functor derivative* [1] for an analytic functor F, which models *one-hole contexts* over F. Consider the collection of functions $\bigsqcup_{u \in U} X^{U \setminus \{u\}}$, which can be seen as the collection of partial functions from U to X that are undefined precisely at one element. A group H of permutations on U acts on $\bigsqcup_{u \in U} X^{U \setminus \{u\}}$ by $\sigma \cdot (u, \phi) = (\sigma(u), \phi \circ (\sigma^{-1}|_{U \setminus \{\sigma(u)\}}))$, for $u \in U$, $\phi : U \setminus \{u\} \to X$. The orbit of an element $(u, \phi) \in \bigsqcup_{u \in U} X^{U \setminus \{u\}}$ is denoted by $[u, \phi]_{H_i}$. The functor derivative of F is the functor $F'(X) = \bigsqcup_{i \in I} \left(\bigsqcup_{u \in U_i} X^{U_i \setminus \{u\}} \right)/H_i$. Elements of $F'(X)$ are of the form $(i, [u, \phi]_{H_i})$ and are called one-hole contexts, because one position is empty. An element $x \in X$ can be "plugged" into a context $[u, \phi]_{H_i} \in F'X$, resulting in $[\phi \cup \{(u, x)\}]_{H_i} \in FX$. Formally, define the *context plug-in* natural transformation $\triangleright : F' \times Id \Rightarrow F$ by $\triangleright_X([u, \phi]_{H_i}, x) := [\phi \cup \{(u, x)\}]_{H_i}$.

Proposition 2.1 *The plug-in is weakly cartesian, i.e., every naturality square of $\triangleright$ is a weak pullback.*

We often use the following notational convention: given a set X, write $x \in X$, $\bar{x} \in FX$ and $\bar{x}' \in F'X$.

Analytic functors and their derivatives satisfy the conditions for the existence of a base. Concretely, their base is given by $\mathrm{Base}_F([\phi]_{H_i}) = \mathrm{Im}(\phi)$ and $\mathrm{Base}_{F'}([u, \phi]) = \mathrm{Im}(\phi)$ for $\phi \in FX$, $[u, \phi]_{H_i} \in F'X$. We have the property $\mathrm{Base}_F(\triangleright_X(\bar{x}', x)) = \mathrm{Base}_{F'}(\bar{x}') \cup \{x\}$, for $\bar{x}' \in F'X$ and $x \in X$. Moreover, if $\bar{x} \in FX$ and $x \in \mathrm{Base}_F(\bar{x})$, there exists a (not necessarily unique) $\bar{x}' \in F'X$ with $\triangleright_X(\bar{x}', x) = \bar{x}$ (see [5] for details). Analytic functors also preserve weak pullbacks.

We introduce a new *context decomposition* natural transformation. Intuitively, context decomposition $\square : F \Rightarrow F(F' \times Id)$ does the opposite of context plug-in: it gives all possible ways to split $\bar{x} \in FX$ into a context in $\bar{x}' \in F'X$ and an element in $x \in X$. Moreover, it organises all decompositions $(\bar{x}', x)$ of $\bar{x}$ into an F-structure, based on the position of the context hole. For each such decomposition $(\bar{x}', x)$, think of $\bar{x}'$ as the context of siblings of x in $\bar{x}$. This will be essential in Definition 5.1 (the algebraic automaton).

4

Definition 2.2 Given an analytic functor $F = \bigsqcup_{i \in I} X^{U_i}/H_i$, define the *context decomposition* natural transformation $\boxdot : F \Rightarrow F(F' \times Id)$ as follows:

$$\boxdot_X(i, [\phi]_{H_i}) := (i, [\psi]_{H_i}), \quad \text{where} \quad \psi : U_i \to F'X \times X, \quad \psi(u) := ((i, [u, \phi \setminus \{\langle u, \phi(u) \rangle\}]_{H_i}), \phi(u)).$$

Example 2.3 Take $F(X) = \Sigma \times X^3$ where $\Sigma = \{a, b\}$. Then $F'(X) = \Sigma \times 3 \times X^2$, and for $X = \{x_0, x_1, x_2\}$ and $(a, x_0, x_1, x_2) \in FX$, we have:

$$\boxdot_X(a, x_0, x_1, x_2) = (a, ((a, 0, x_1, x_2), x_0), ((a, 1, x_0, x_2), x_1), ((a, 2, x_0, x_1), x_2)).$$

Example 2.4 Take $F = \mathcal{B}_3$, the bag functor where the bag size is bounded by 3, i.e., $F = \bigsqcup_{n \leq 3} X^n/H$ where H is the symmetric group on n. Then $F' = \mathcal{B}_2$. We use the notation $\{\dots\}_b$ for bags. For $X = \{x_0, x_1, x_2\}$, we have:

$$\boxdot_X(\{x_0, x_0, x_1\}_b) = \{(\{x_0, x_1\}_b, x_0), (\{x_0, x_1\}_b, x_0), (\{x_0, x_0\}_b, x_1)\}_b.$$

In order to avoid working with the concrete definition of $\boxdot$, we identify its key abstract properties.

Lemma 2.5 *Context decomposition* $\boxdot$ *satisfies:*

(i) $F\mathsf{pr}_2 \circ \boxdot_X = id$ *(see Figure 2a);*

(ii) *for every element* $\upsilon = (\bar{y}', y) \in F'(F'X \times X) \times (F'X \times X)$ *with* $\rhd_{F'X \times X}(\upsilon) \in \boxdot_X[FX]$, *we have* $F'\mathsf{pr}_2(\bar{y}') = \mathsf{pr}_1(y)$ *(see Figure 2b);*

(iii) *if* $\bar{x} \in FX$ *and* $(\bar{x}', x) \in \text{Base}_F(\boxdot_X(\bar{x}))$, *then* $\rhd_X(\bar{x}', x) = \bar{x}$.

(a) Property (i)

(b) Property (ii)

Fig. 2. Diagrams for Lemma 2.5.

Property (i) completely describes the content of the *Id*-component of $\boxdot_X(\bar{x}) \in F(F' \times Id)$. Together with property (i), property (ii) completely describes the F'-component. Thus these two properties can be taken as an abstract, equivalent definition of $\boxdot$. Property (iii) follows from (i) and (ii) and it conveys our intuitive understanding that $\boxdot$ decomposes $\bar{x} \in FX$ into pairs of an element x and its siblings $\bar{x}'$.

We apply the concept of *relation lifting* [13] for analytic functors. Specifically, we will use the lifting $\bar{\in}$ of the "element of" relation $\in$. Given a set X, $p \in FX$ and $q \in F(\mathcal{P}(X))$, we have $p \bar{\in} q$ if there exists $r \in F(\in)$ such that $F\mathsf{pr}_1(r) = p$ and $F\mathsf{pr}_2(r) = q$ (we write pr_1 and pr_2 for product projections, also in_1 and in_2 for coproduct injections). Informally, $p \bar{\in} q$ means "p and q have matching indices in I and p is position-wise contained in q". The parameters F and X, on which $\bar{\in}$ depends, are left implicit and understood from the context.

Assumption. For the rest of the paper, we fix an analytic functor $FX = \bigsqcup_{i \in I} X^{U_i}/H_i$ where I is finite. This ensures that F and F' preserve finite sets.

2.4 Thin Coalgebras

We are interested in running F-automata on a subclass of F-coalgebras called *thin F-coalgebras* [5]. Thin coalgebras generalise the notion of *thin tree* [16] to the level of coalgebras. They are defined as those F-coalgebras for which every state is the starting point of only countably many infinite paths. More precisely, given a F-coalgebra (X, ξ) and $x \in X$ with $\xi(x) = (i, [\phi]_{H_i})$, we say that an element $x_1 \in \text{Base}_F(i, [\phi]_{H_i}) = \text{Im}(\phi)$ is a *successor* of x with multiplicity $|\phi^{-1}(x_1)|$. Thus the successor relation on (X, ξ) defines a multigraph, with multiplicities corresponding to multiple parallel edges. A state $x \in X$ is *thin* if there are

$$\begin{array}{ccccc}
(F+G)A & \xrightarrow{(F+G)\llbracket-\rrbracket} & (F+G)Z^{\flat} & \longrightarrow & (F+G)Z \\
{\scriptstyle\alpha}\downarrow & & {\scriptstyle\beta^{\flat}}\downarrow & & \downarrow{\scriptstyle\beta} \\
A & \xrightarrow{\quad\llbracket-\rrbracket\quad} & Z^{\flat} & \lhook\joinrel\longrightarrow & Z \\
& & {\scriptstyle\zeta^{\flat}}\downarrow & & \downarrow{\scriptstyle\zeta} \\
& & FZ^{\flat} & \longrightarrow & FZ
\end{array}$$

Fig. 3. Algebra and coalgebra on $Z^{\flat}$.

only countably many infinite paths starting from x in this multigraph. A (pointed) coalgebra is *thin* if all its states are thin.

Behaviours of thin coalgebras can be characterised algebraically via *coherent $(F+G)$-algebras*. Define the functor $G(X) \coloneqq (F'X)^{\omega}$, mapping X to the set of streams of contexts over X. An $(F+G)$-algebra is of the form (C, γ), with $\gamma = [\gamma_0, \gamma_1]$, where $\gamma_0 : FX \to X$ is an F-algebra structure and $\gamma_1 : GX \to X$ is a G-algebra structure. An $(F+G)$-algebra (C, γ) is *coherent* if it satisfies the equation $\gamma_1 = \gamma_0 \circ \rhd_C \circ \langle id, \gamma_1 \rangle \circ \langle hd, tl \rangle$, where hd stands for stream head and tl stands for stream tail. Roughly, the equation says "evaluating a stream with γ_1 is equal to evaluating the stream tail, plugged into the stream head, with γ_0". [5, Corollary VII.6] shows that the initial coherent $(F+G)$-algebra is isomorphic to the collection of behaviours of thin coalgebras.

The initial coherent $(F+G)$-algebra is given concretely as follows. Fix an initial $(F+G)$-algebra $(A, \alpha = [\alpha_0, \alpha_1])$ and a final F-coalgebra (Z, ζ) (their existence is proven in [5]). There exists a natural way to interpret terms $a \in A$ in Z. Informally speaking, for $\bar{a} \in FA$, $a = \alpha_0(\bar{a})$ is interpreted as a state with successors $\bar{a}$; for $(\bar{a}'_n)_{n\in\omega} \in GA$, $a = \alpha_1((\bar{a}')_{n\in\omega})$ is interpreted by successively plugging all contexts $(\bar{a}'_n)_{n\in\omega}$ into each other, i.e., plugging $\bar{a}'_1$ into $\bar{a}'_0$, $\bar{a}'_2$ into $\bar{a}'_1$, $\bar{a}'_3$ into $\bar{a}'_2$ and so on. This is formalised by defining a suitable $(F+G)$-algebra structure $\beta = [\beta_0, \beta_1]$ on Z and taking the *interpretation map* $\llbracket-\rrbracket : (A, \alpha) \to (Z, \beta)$ to be the unique map obtained by initiality of (A, α). By taking the image $Z^{\flat} \subseteq Z$ [5] of the interpretation map, one obtains both an $(F+G)$-subalgebra $(Z^{\flat}, \beta^{\flat})$ of (Z, β) and an F-subcoalgebra $(Z^{\flat}, \zeta^{\flat})$ of (Z, ζ). Figure 3 gives a visual summary. We have that $(Z^{\flat}, \beta^{\flat})$ is an initial coherent $(F+G)$-algebra, i.e., for every coherent $(F+G)$-algebra (C, γ), there exists a unique $(F+G)$-algebra morphism $cev_{(C,\gamma)} : (Z^{\flat}, \beta^{\flat}) \to (C, \gamma)$. Moreover, $(Z^{\flat}, \zeta^{\flat})$ is a final thin coalgebra, meaning that for every thin coalgebra (X, ξ), there exists a unique F-coalgebra morphism $tbeh_{(X,\xi)} : (X, \xi) \to (Z^{\flat}, \zeta^{\flat})$. In other words, $(Z^{\flat}, \zeta^{\flat})$ is the subcoalgebra of all *thin behaviours*, i.e., behaviours of thin coalgebras. Furthermore, $Z^{\flat}$ is isomorphic to the collection of *normal terms* [5, Section V]: each $z \in Z^{\flat}$ has a canonical *normal representative* $a \in A$ with $\llbracket a \rrbracket = z$. We have two useful properties connecting $\beta^{\flat}$ and $\zeta^{\flat}$:

$$\zeta^{\flat} = (\beta_0^{\flat})^{-1}, \tag{1}$$

$$(z_m)_{m\in\omega} \in (Z^{\flat})^{\omega}, (\bar{z}'_m)_{m>0} \in (F'Z^{\flat})^{\omega}, \forall n \in \omega : \rhd_{Z^{\flat}}(\bar{z}'_{m+1}, z_{m+1}) = \zeta^{\flat}(z_m) \implies \beta_1^{\flat}((\bar{z}'_m)_{n>0}) = z_0. \tag{2}$$

In the present work, we are interested in the language recognition aspect of $(F+G)$-algebras. Given a coherent $(F+G)$-algebra (C, γ) and $U \subseteq C$, the *language* of the triple (C, γ, U) is defined as $L(C, \gamma, U) \coloneqq cev_{(C,\gamma)}^{-1}(U) \subseteq Z^{\flat}$. Hence coherent algebras recognise languages of thin behaviours, similarly to how ω-semigroups recognise languages of infinite words. We refer to such a triple (C, γ, U) as a coherent $(F+G)$-algebra with a *recognising set*.

3 Runs and Unambiguity of F-Coalgebra Automata

In this section, we present *F-coalgebra automata* (for brevity, *F*-automata), which were studied in [12] as automata accepting F-coalgebras. We define acceptance of F-automata via the notion of *run*, in contrast

---

[5] The superscript $\flat$ is pronounced as "thin". The letter *thorn* $\flat$ denotes a voiceless dental fricative (the first sound in "thin") in Old English.

with [12], which defines acceptance via parity games. Our reason for introducing runs is to be able to define *unambiguous F-automata*. While the two definitions of acceptance (via runs and via parity games) appear to coincide, we do not show it in this paper, as we work exclusively with runs. We note that a similar definition of F-automaton runs and unambiguity is given in [6], but only for polynomial functors F.

Definition 3.1 An *F-automaton* is a quadruple $\mathcal{A} = (Q, \delta, Q_I, \text{Acc})$ where Q is a finite set of *states*, $\delta \colon Q \to (\mathcal{P} \circ F)(Q)$ is a *transition function*. $Q_I \subseteq Q$ is a set of *initial states*, and $\text{Acc} \subseteq Q^\omega$ is an *acceptance condition*.

According to the above definition, F-automata are, in general, nondeterministic, i.e., every state $q \in Q$ has an arbitrary set $\delta(q)$ of transitions and there are multiple initial states Q_I. We do not place any conditions of the acceptance condition; instead, we distinguish the following types of acceptance conditions.

Definition 3.2 Let $\mathcal{A} = (Q, \delta, Q_I, \text{Acc})$ be an F-automaton. We call Acc:

- *parity* if there exists a map $\Omega : Q \to \mathbb{N}$ such that $(q_n)_{n \in \omega} \in \text{Acc}$ if and only if $\limsup_{n \in \omega} \Omega(q_n)$ is even;
- *ω-regular* if Acc is an ω-regular language over the alphabet Q;
- *prefix-agnostic* if for all $x \in Q^\omega$, $w \in Q^*$: $wx \in \text{Acc}$ if and only if $x \in \text{Acc}$.

F-automata with a parity acceptance condition are known as *parity F-automata* and we write them as a tuple (Q, δ, Q_I, Ω), with Ω instead of Acc. By taking the polynomial functor $F(X) = \Sigma \times X$, for some alphabet Σ, we obtain nondeterministic parity word automata.

Since parity word automata recognise ω-regular languages, one can see that every parity condition is also ω-regular. Conversely, every F-automaton with ω-regular acceptance can be turned into an equivalent parity F-automaton via the *wreath product* construction [12, Theorem 4.4]. Parity conditions are also prefix-agnostic, but automata with the prefix-agnostic conditions turn out to be strictly more expressive, as shown below.

Example 3.3 Consider the functor $F(X) = \Sigma \times X$, for $\Sigma = \{a, b\}$, whose derivative is $F'(X) = \Sigma$. Define the (word) F-automaton $\mathcal{A} = (Q, \delta, Q_I, \text{Acc})$ with $Q \coloneqq \{q_a, q_b\}$, $\delta(q) \coloneqq \{(a, q_a), (b, q_b)\}$ for all $q \in Q$, $Q_I \coloneqq Q$. Let $\text{Acc} \subseteq Q^\omega$ consist of those infinite words that contain infinitely many q_b's and unboundedly many consecutive q_a's, i.e., for every natural number n, the word contains n-many consecutive q_a's. One readily sees that Acc is prefix-agnostic and $\mathcal{A}$ accepts (in the classical sense) the language $L \coloneqq \{a_0 a_1 \ldots \mid q_{a_0} q_{a_1} \ldots \in \text{Acc}\}$. However, L is not ω-regular. This is because every non-empty ω-regular language contains an ultimately periodic word, i.e., a word of the form wu^ω, while L contains no such words. This example shows that automata with prefix-agnostic acceptance are more expressive than parity automata.

Next, we define F-automaton runs and unambiguity, thereby generalising the definitions in [6] from polynomial functors to arbitrary analytic functors. Below we write Δ_Y for the constant functor sending every set to the set Y and every function to id_Y.

Definition 3.4 Let $\mathcal{A} = (Q, \delta, Q_I, \text{Acc})$ be an F-automaton and (X, ξ, x_I) be a pointed F-coalgebra. A *pre-run* of $\mathcal{A}$ on (X, ξ, x_I) is a reachable pointed $(F \times \Delta_X \times \Delta_Q)$-coalgebra $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ satisfying:

(i) $\rho_X : (R, \rho_F, r_I) \to (X, \xi, x_I)$ is a pointed F-coalgebra morphism;

(ii) $(F\rho_Q \circ \rho_F)(r) \in (\delta \circ \rho_Q)(r)$ for all $r \in R$;

(iii) $(\rho_Q(r_n))_{n \in \omega} \in \text{Acc}$ for all $(r_n)_{n \in \omega} \in R^\omega$ with $r_0 = r_I$ and $\forall n(r_{n+1} \in \text{Base}_F(\rho_F(r_n)))$.

We define a *run* as a pre-run (R, ρ, r_I) for which (R, ρ) is a subcoalgebra of the final $(F \times \Delta_X \times \Delta_Q)$-coalgebra. A (pre-)run is *accepting* if $\rho_Q(r_I) \in Q_I$. We say that $\mathcal{A}$ *accepts* (X, ξ, x_I) if there exists an accepting run of $\mathcal{A}$ on (X, ξ, x_I).

Definition 3.5 Let $\mathcal{A}$ be an F-automaton and (X, ξ, x_I) be a pointed F-coalgebra. We say that $\mathcal{A}$ is *unambiguous* on (X, ξ, x_I) if $\mathcal{A}$ has at most one accepting run on (X, ξ, x_I).

A pre-run of $\mathcal{A}$ on (X, ξ, x_I) represents an execution of $\mathcal{A}$ on the structure of (X, ξ, x_I). The pre-run is a span $X \xleftarrow{\rho_X} R \xrightarrow{\rho_Q} Q$ where R is equipped with an F-coalgebra structure $\rho_F : R \to F(R)$. Property
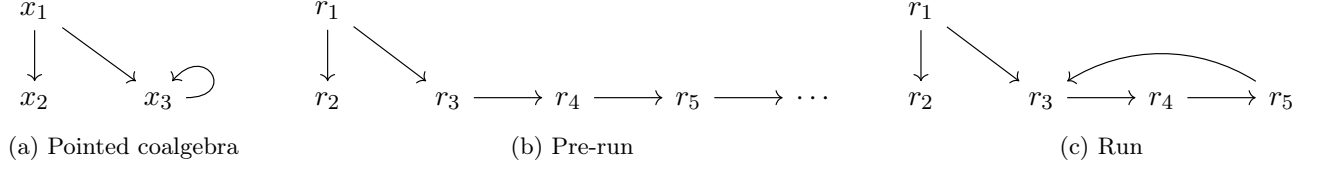
(a) Pointed coalgebra   (b) Pre-run   (c) Run

Fig. 4. Examples of (pre-)runs

(i) says that $\rho_X \colon R \to X$ respects the coalgebra structure of X; property (ii) says that $\rho_Q \colon R \to Q$ respects the automaton transitions; and property (iii) says that ρ_Q respects the automaton acceptance condition. We note that, classically, automaton runs need not respect the acceptance condition (they are called *final* if they do) but in this paper, all (pre-)runs are required to be final. Runs have the additional property that behaviourally equivalent elements are identified. This is essential for the definition of unambiguous automata, where we count the number of accepting runs of an automaton modulo behavioural equivalence. Intuitively, runs are as close as possible to being a relation between X and Q. However, there could be elements $r_1 \neq r_2 \in R$ of a run R with $\rho_X(r_1) = \rho_X(r_2)$ and $\rho_Q(r_1) = \rho_Q(r_2)$, since r_1 and r_2 can still differ on ρ_F.

Example 3.6 Consider the polynomial functor $F = X^2 + X + 1$. Figure 4a depicts a pointed F-coalgebra (X, ξ, x_1) with $\xi(x_1) = (x_2, x_3)$, $\xi(x_2) = ()$ and $\xi(x_3) = (x_3)$. Consider the F-automaton $\mathcal{A} = (Q, \delta, Q_I, \mathrm{Acc})$ with $Q = \{q_1, q_2, q_3, q_4\}$, $Q_I := \{q_1\}$, $\delta(q_1) = \{(q_2, q_3)\}$, $\delta(q_2) = \{()\}$, $\delta(q_3) = \{(q_3), (q_4)\}$, $\delta(q_4) = \{(q_3)\}$ and Acc containing all $(q_n)_{n\in\omega} \in Q^\omega$ where $q_n = q_3$ for infinitely many $n \in \omega$. Figure 4b depicts the F-structure ρ_F of an accepting pre-run $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_1)$, with ρ_X and ρ_Q given by:

$$\rho_X \colon \quad r_1 \mapsto x_1, \quad r_2 \mapsto x_2, \quad \{r_3, r_4, \dots\} \mapsto x_3,$$
$$\rho_Q \colon \quad r_1 \mapsto q_1, \quad r_2 \mapsto q_2, \quad \{r_3, r_4, r_6, r_7, \dots, r_{3n}, r_{3n+1}, \dots\} \mapsto q_3, \quad \{r_5, r_8, \dots, r_{3n+2}, \dots\} \mapsto q_4.$$

Note that (R, ρ, r_1) is not (isomorphic to) a run, because r_3 and r_6 are behaviourally equivalent. Figure 4c shows the F-structure of another accepting pre-run $(R', \rho' = \langle \rho'_F, \rho'_X, \rho'_Q \rangle, r_1)$, with ρ'_X and ρ'_Q given by:

$$\rho'_X \colon \quad r_1 \mapsto x_1, \quad r_2 \mapsto x_2, \quad \{r_3, r_4, r_5\} \mapsto x_3,$$
$$\rho'_Q \colon \quad r_1 \mapsto q_1, \quad r_2 \mapsto q_2, \quad \{r_3, r_4\} \mapsto q_3, \quad r_5 \mapsto q_4.$$

One can check that (R', ρ', r_1) is isomorphic to a run, since no two elements of R' are behaviourally equivalent. Moreover, by merging r_3 and r_4 (i.e., setting $\rho'_F(r_3) = r_5$ and dropping r_4), we obtain another accepting run. Therefore $\mathcal{A}$ is not unambiguous on (X, ξ, x_1).

Below we state two basic properties of (pre-)runs.

Lemma 3.7 *Let $\mathcal{A}$ be an F-automaton and (X, ξ, x_I) be a pointed F-coalgebra. If (R, ρ, r_I) is an accepting pre-run of $\mathcal{A}$ on (X, ξ, x_I), then its image in the final $(F \times \Delta_X \times \Delta_Q)$-coalgebra is an accepting run. Hence $\mathcal{A}$ accepts (X, ξ, x_I) if and only if $\mathcal{A}$ has an accepting pre-run on (X, ξ, x_I).*

Proposition 3.8 *Let $\mathcal{A}$ be an F-automaton and $f \colon (X, \xi, x_I) \to (X', \xi', x'_I)$ be a pointed F-coalgebra morphism. Then $\mathcal{A}$ accepts (X, ξ, x_I) if and only if $\mathcal{A}$ accepts (X', ξ', x'_I).*

We conclude this section with a strengthening of [12, Theorem 4.4]: every automaton with ω-regular acceptance can be transformed into an equivalent parity automaton in an unambiguity-preserving way.

Proposition 3.9 *Every F-automaton $\mathcal{A}$ with ω-regular acceptance can be transformed into a parity F-automaton $\mathcal{A}'$ accepting the same coalgebras. Moreover, for every pointed F-coalgebra (X, ξ, x_I), if $\mathcal{A}$ is unambiguous on (X, ξ, x_I), then $\mathcal{A}'$ is unambiguous on (X, ξ, x_I).*

4 From Automata to Algebras

In this section, we describe the first central construction of the paper: the automaton algebra. Given an F-automaton with prefix-agnostic acceptance, we construct a finite coherent $(F + G)$-algebra with a

recognising set, which recognises the language consisting of the thin behaviours that are accepted by the F-automaton. Our construction is inspired by the construction of the *thin algebra of an automaton* for binary trees [16, Section 6.2.1].

4.1 The Automaton Algebra

Given an F-automaton $\mathcal{A}$, the elements of the automaton algebra $\mathsf{C}_{\mathcal{A}}$ will be sets of automaton states. The algebra structure of $\mathsf{C}_{\mathcal{A}}$ is defined in order to obtain the following property: if $z \in Z^{\flat}$, then $cev_{\mathsf{C}_{\mathcal{A}}}(z)$ is the set of those states q such that $\mathcal{A}$ has a run of $(Z^{\flat}, \zeta^{\flat}, z)$, starting at q (recall $cev_{\mathsf{C}_{\mathcal{A}}} : (Z^{\flat}, \beta^{\flat}) \to \mathsf{C}_{\mathcal{A}}$).

Definition 4.1 Let $\mathcal{A} = (Q, \delta, Q_I, \mathrm{Acc})$ be an F-automaton with prefix-agnostic Acc. Define the *automaton algebra* $\mathsf{C}_{\mathcal{A}} := (C, [\gamma_0, \gamma_1], U)$ of $\mathcal{A}$ as follows.

- $C := \mathcal{P}(Q)$;
- for all $\bar{c} \in FC$: $\gamma_0(\bar{c}) := \{q \in Q \mid \exists \bar{q} \in FQ(\bar{q} \,\bar{\in}\, \bar{c} \wedge \bar{q} \in \delta(q))\}$;
- for all $(\bar{c}'_n)_{n \in \omega} \in (F'C)^{\omega}$:

$$\gamma_1((\bar{c}'_n)_{n \in \omega}) := \{q_0 \in Q \mid \exists (q_n)_{n \in \omega} \in \mathrm{Acc}, (\bar{q}'_n)_{n \in \omega} \in (F'Q)^{\omega} : \forall n \in \omega(\bar{q}'_n \,\bar{\in}\, \bar{c}'_n),$$
$$\forall n \in \omega(\triangleright_Q(\bar{q}'_n, q_{n+1}) \in \delta(q_n))\};$$

- $U := \{c \in C \mid c \cap Q_I \neq \emptyset\}$.

For simplicity, consider a polynomial functor F. In the definition of γ_0, $\gamma_0(\bar{c})$ consists of those states q, for which there exists a transition $\bar{q} \in \delta(q)$ such that each component in the tuple $\bar{q}$ is an element of the corresponding component of $\bar{c}$. This corresponds to the fact that a thin behaviour z is accepted by $\mathcal{A}$, starting at q, precisely when there exists a transition $\bar{q} \in \delta(q)$ such that, for all i, $\mathcal{A}$ accepts the i-th successor of z, starting at the i-th component of $\bar{q}$. Here it is essential to assume Acc is prefix-agnostic, so that for all $p \in \mathrm{Base}_F(\bar{q})$ and $x \in Q^{\omega}$, we have $qpx \in \mathrm{Acc}$ if and only if $px \in \mathrm{Acc}$.

Similarly, $\gamma_1((\bar{c}'_n)_{n \in \omega})$ consists of states q such that we can choose a context $\bar{q}'_n$ of states for every context $\bar{c}'_n$, and a sequence of states $(q_n)_{n \geq 1} \in \mathrm{Acc}$ to fill the consecutive holes in these contexts. Again, we use the prefix-agnostic assumption, so that for all $p \in \mathrm{Base}_{F'}(\bar{q}'_n)$ and $x \in Q^{\omega}$, we have $q_0 q_1 \ldots q_n px \in \mathrm{Acc}$ if and only if $px \in \mathrm{Acc}$.

For the recognising set U, we take those sets of states c that contain at least one accepting state, so that $cev_{\mathsf{C}_{\mathcal{A}}}^{-1}(U)$ contains the thin behaviours accepted by $\mathcal{A}$.

Theorem 4.3 below connects acceptance of an automaton $\mathcal{A}$ with the language $L(\mathsf{C}_{\mathcal{A}})$ of its automaton algebra. It uses the key property that the automaton algebra is coherent.

Lemma 4.2 *For all F-automata $\mathcal{A}$ with prefix-agnostic acceptance, the automaton algebra $\mathsf{C}_{\mathcal{A}}$ is coherent.*

Theorem 4.3 *Let $\mathcal{A}$ be an F-automaton with prefix-agnostic acceptance and let (X, ξ, x_I) be a thin pointed F-coalgebra. Then $\mathcal{A}$ accepts (X, ξ, x_I) if and only if $tbeh_{(X,\xi)}(x_I) \in L(\mathsf{C}_{\mathcal{A}})$.*

Proof (Sketch) Suppose $\mathcal{A} = (Q, \delta, Q_I, \mathrm{Acc})$ and $\mathsf{C}_{\mathcal{A}} = (C, \gamma = [\gamma_0, \gamma_1], U)$. Define:

$$f : Z^{\flat} \to C,$$
$$z \mapsto \{q \in Q \mid \text{there exists a run of } \mathcal{A} \text{ on } (Z^{\flat}, \zeta^{\flat}, z), \text{ starting from } q\}.$$

One can show that $f : (Z^{\flat}, \beta^{\flat}) \to (C, \gamma)$ is an $(F + G)$-algebra morphism. Since $(Z^{\flat}, \beta^{\flat})$ is an initial coherent $(F + G)$-algebra, this implies $f = cev_{(C,\gamma)}$. Now $\mathcal{A}$ accepts (X, ξ, x_I) if and only if $\mathcal{A}$ accepts $(Z^{\flat}, \zeta^{\flat}, tbeh_{(X,\xi)}(x_I))$ (by Proposition 3.8) if and only if $(f \circ tbeh_{(X,\xi)})(x) \cap Q_I \neq \emptyset$ if and only if $(cev_{(C,\gamma)} \circ tbeh_{(X,\xi)})(x) \in U$ if and only if $tbeh_{(X,\xi)}(x_I) \in L(\mathsf{C}_{\mathcal{A}})$. $\qquad\square$

Example 4.4 Let $\mathcal{A}$ be the automaton from Example 3.3. Its automaton algebra $\mathsf{C}_{\mathcal{A}} = (C, [\gamma_0, \gamma_1], U)$ has a carrier $C = \{\emptyset, \{q_a\}, \{q_b\}, \{q_a, q_b\}\}$. The F-operation is given by $\gamma_0(\sigma, c) = c$, for $\sigma \in \Sigma$ and $c \in C$. For the G-operation, for every $(\sigma_n)_{n \in \omega} \in GC \cong \Sigma^{\omega}$, we have that $\gamma_1((\sigma_n)_{n \in \omega})$ equals $\{q_a, q_b\}$, if $(\sigma_n)_{n \in \omega} \in L$,

and $\emptyset$, otherwise. For the recognising set, we have $U = \{\{q_a\}, \{q_b\}, \{q_a, q_b\}\}$. If we take $Z = \Sigma^\omega$ (the final coalgebra of streams over Σ), we get $L(\mathsf{C}_\mathcal{A}) = L \subseteq Z^\flat = Z$.

4.2 Rational Algebras

Example 4.4 showed that there exist finite coherent $(F+G)$-algebras whose language cannot be characterised by parity F-automata. A finite coherent $(F + G)$-algebra (C, γ) partitions $Z^\flat$ into finitely many classes $\{cev_{(C,\gamma)}^{-1}(c) \mid c \in C\}$. In order to retain the connection to parity F-automata, in Definition 4.5 we equip (C, γ, U) with additional structure so that it also partitions into finitely many classes the set $(F'Z^\flat)^+$ of finite sequences of contexts over $Z^\flat$. Intuitively, a sequence of n contexts is viewed as the "nested context" obtained by plugging the sequence together, so that the hole is at depth n (whereas in our usual contexts the hole is at depth 1). The partition of $(F'Z^\flat)^+$ is to satisfy the following property: if $(\widetilde{c}_n)_{n\in\omega}, (\widetilde{d}_n)_{n\in\omega} \in ((F'Z^\flat)^+)^\omega$ and for all $n \in \omega$, $\widetilde{c}_n$ and $\widetilde{d}_n$ are in the same class, then $\gamma_1(\widetilde{c}_0\widetilde{c}_1 \dots) \in L(C, \gamma, U)$ if and only if $\gamma_1(\widetilde{d}_0\widetilde{d}_1 \dots) \in L(C, \gamma, U)$. Note that in Example 4.4 it is impossible to find a finite partition of $(F'Z^\flat)^+ = \Sigma^+$ with this property. In order to guarantee the property, below we identify the following subclass of finite coherent $(F + G)$-algebras.

Definition 4.5 Let $(C, \gamma = [\gamma_0, \gamma_1])$, be a finite coherent $(F + G)$-algebra and $\Sigma := F'C$. We call (C, γ) *rational* if there exists a finite ω-semigroup $(\widetilde{C}, \mathrm{Im}(\gamma_1))$ and a map $\gamma_2 : \Sigma^+ \to \widetilde{C}$ such that $(\gamma_2, \gamma_1) : (\Sigma^+, \Sigma^\omega) \to (\widetilde{C}, \mathrm{Im}(\gamma_1))$ is an ω-semigroup homomorphism.

In the above definition, the map $\gamma_2 : \Sigma^+ \to \widetilde{C}$ partitions the set of finite sequences of contexts (i.e., the nested contexts) into finitely many classes $\widetilde{C}$.

Note that for a functor $F = \Sigma_0 + \Sigma_2 \times X \times X$, where Σ_0 and Σ_2 are alphabets, rational $(F+G)$-algebras essentially coincide with thin algebras [16]. Thin algebras contain two sorts: a sort for trees (in rational algebras, this is the domain C) and a sort for contexts (in rational algebras, this is the set $\widetilde{C}$). Hence rational $(F + G)$-algebras can be seen as a natural generalisation of thin algebras to analytic functors.

We will see in Section 6 that languages of rational $(F + G)$-algebras can be characterised by parity F-automata. For now, we only show that parity F-automata give rise to rational $(F + G)$-algebras.

Proposition 4.6 *For all F-automata $\mathcal{A}$ with parity acceptance, the automaton algebra $\mathsf{C}_\mathcal{A}$ is rational.*

Proof (Sketch) The construction generalises [16, Section 6.2.1]. Let $\mathcal{A} = (Q, \delta, Q_I, \Omega)$ be a parity F-automaton, and $\mathsf{C}_\mathcal{A} = (C, \gamma = [\gamma_0, \gamma_1], U)$. Define a two-sorted algebra $(\widehat{C}, C)$ by:

$$\widehat{C} := \mathcal{P}(Q \times Q \times \mathrm{Im}(\Omega)),$$
$$\widehat{c}_1 \cdot \widehat{c}_2 := \{(q, q_2, \max\{m_1, m_2\}) \mid \exists q_1 \in Q : (q, q_1, m_1) \in \widehat{c}_1 \wedge (q_1, q_2, m_2) \in \widehat{c}_2\},$$
$$\widehat{c} \times c := \{q \mid \exists q_1 \in c, m \in \omega : (q, q_1, m) \in \widehat{c}\},$$
$$\Pi((\widehat{c}_n)_{n\in\omega}) := \{q_0 \mid \exists (q_n)_{n\in\omega} \in Q^\omega, (m_n)_{n\in\omega} \in \mathbb{N}^\omega : \forall n \in \omega((q_n, q_{n+1}, m_n) \in \widehat{c}_n) \wedge \limsup_{n\in\omega} m_n \text{ is even}\}.$$

We define the map $\gamma_2 : (F'C)^+ \to \widehat{C}$ by specifying its restriction to the set of generators $F'C$ of the freely generated semigroup $(F'C)^+$. For $\bar{c}' \in F'C$, we set:

$$\gamma_2(\bar{c}') := \{(q, q_1, \max\{\Omega(q), \Omega(q_1)\}) \mid \exists \bar{q}' \in F'Q(\bar{q}' \bar{\in} \bar{c}' \wedge \triangleright_Q(\bar{q}', q_1) \in \delta(q))\}.$$

One can show that $(\mathrm{Im}(\gamma_2), \mathrm{Im}(\gamma_1))$ is an ω-semigroup and (γ_2, γ_1) is a homomorphism. $\square$

5 From Algebras to Automata

In this section, we show how to construct from a finite coherent algebra its *algebraic automaton*. The context decomposition transformation from Definition 2.2 is instrumental in defining the transition structure of this automaton. The key result here is that the algebraic automaton is unambiguous on thin coalgebras. We

proceed as follows: we introduce the algebraic automaton, develop the key technical notion of *marking* and use it to show that, when restricting to thin F-coalgebras, the algebraic automaton is unambiguous and accepts the same language as the starting algebra.

5.1 The Algebraic Automaton

Given a finite coherent $(F + G)$-algebra $(C, [\gamma_0, \gamma_1], U)$ with a recognising set, we aim to construct an equivalent unambiguous automaton. We draw inspiration from the construction in [16, Section 7.2.1] for binary trees. The idea is that each state q in the algebraic automaton encodes an element $c \in C$ in such a way that the algebraic automaton accepts, starting at state q, those pointed coalgebras $(Z^\flat, \zeta^\flat, z)$ for which $cev_{(C,\gamma)}(z) = c$. A run of the algebraic automaton labels behaviours $z \in Z^\flat$ with algebra elements $c \in C$. The transitions of the automaton are to ensure that if z is labelled with $c \in C$ and $\zeta^\flat(z) \in FZ^\flat$ is labelled with $\bar{c} \in FC$, then $\gamma_0(\bar{c}) = c$. The acceptance condition is to ensure that for every infinite path $(z, z_1, z_2 \dots)$, if z is labelled with c, and for all $n \geq 1$, the context in $F'Z^\flat$ consisting of the siblings of z_n is labelled with $\bar{c}'_n \in F'C$, then $\gamma_1((\bar{c}'_n)_{n \geq 1}) = c$. In order to realise the latter requirement, a state q of the automaton must encode both a label $c \in C$ and a context $\bar{c} \in F'C$ for the label of its sibling context. Since the root of a pointed F-coalgebra does not have any siblings, we need additional states that only encode a label in C – these states occur only in the root of the run.

   Recall that the notion "context of siblings" can be expressed formally using the context decomposition operator $\boxdot : F \Rightarrow F(F' \times Id)$ from Definition 2.2.

Definition 5.1 Let $\mathsf{C} = (C, [\gamma_0, \gamma_1], U)$ be a finite coherent $(F + G)$-algebra with a recognising set. Define the *algebraic automaton* $\mathcal{A}_\mathsf{C} := (Q, \delta, Q_I, \mathrm{Acc})$ as follows:

- $Q := C + F'C \times C$;
- $Q_I := \mathsf{in}_1[U]$;
- $\delta(\mathsf{in}_1(c)) := \{(F\mathsf{in}_2 \circ \boxdot_C)(\bar{c})) \mid \gamma_0(\bar{c}) = c\}$, for $c \in C$, $\delta(\mathsf{in}_2(\bar{c}', c)) := \delta(\mathsf{in}_1(c))$, for $(\bar{c}', c) \in F'C \times C$;
- $\mathrm{Acc} := \{\mathsf{in}_1(c_0) \cdot (\mathsf{in}_2(\bar{c}'_n, c_n))_{n>0} \mid \forall m(c_m = \gamma_1((\bar{c}'_n)_{n>m}))\}$.

   In the above definition of δ, transitions from an automaton state labelled with $c \in C$ cover all possible decompositions of all $\bar{c}$ such that $\gamma_0(\bar{c}) = c$. The algebraic automaton is defined such that it accepts the same thin behaviours as the corresponding coherent algebra. Furthermore, it has precisely one run on each thin coalgebra, thus it is unambiguous. The rest of the section is dedicated to proving these statements.

5.2 Markings

In order to relate pre-runs of the algebraic automaton with the corresponding coherent algebra, we introduce the notion of *marking*, which generalises consistent markings on binary trees [16, Section 7.1].

Definition 5.2 Let $(C, \gamma = [\gamma_0, \gamma_1])$ be a coherent $(F + G)$-algebra and let (X, ξ) be an F-coalgebra. A *marking* of (X, ξ) with (C, γ) is a map $\mu : X \to C$ satisfying:

(i) $\mu : (X, \xi) \to (C, \gamma_0)$ is a F-coalgebra-to-algebra morphism;
(ii) for all $(x_n)_{n \in \omega} \in X^\omega$, $(\bar{x}'_n)_{n>0} \in GX$ with $\forall n(\triangleright_X(\bar{x}'_{n+1}, x_{n+1}) = \xi(x_n))$: $\gamma_1(G\mu((\bar{x}'_n)_{n>0})) = \mu(x_0)$.

   Roughly speaking, property (i) of markings is the algebraic counterpart to property (ii) of pre-runs the algebraic automaton, while property (ii) of markings is the algebraic counterpart to property (iii) of pre-runs. So, intuitively, pre-runs of the algebraic automaton compute a marking. The precise connection between markings and pre-runs of the algebraic automaton is given in the following statement.

Proposition 5.3 *Let* $\mathsf{C} = (C, \gamma = [\gamma_0, \gamma_1], U)$ *be a finite coherent $(F + G)$-algebra with a recognising set, and let (X, ξ, x_I) be a pointed F-coalgebra.*

(i) *If $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ is a pre-run of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I), then $[id, \mathsf{pr}_2] \circ \rho_Q : R \to C$ is a marking of (R, ρ_F) with (C, γ) (see Figure 5a).*
(ii) *If $\mu : X \to C$ is a marking of (X, ξ) with (C, γ), then there exists a pre-run $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I) with $\rho_Q(r_I) \in \mathsf{in}_1[C]$ and $[id, \mathsf{pr}_2] \circ \rho_Q = \mu \circ \rho_X$ (see Figure 5b).*

11

$$R \xrightarrow{\rho_Q} C + (F'C \times C) \xrightarrow{[id, \mathsf{pr}_2]} C$$

(a) Property (i)

(b) Property (ii)

Fig. 5. Diagrams for Proposition 5.3.

Proof (Sketch) (i). It can be verified that $[id, \mathsf{pr}_2] \circ \rho_Q : R \to C$ satisfies the properties of markings, using Lemma 2.5. The proof of property (i) of markings uses property (ii) of the pre-run (R, ρ, r_I), while for property (ii) of markings we use property (iii) of pre-runs.

(ii). We define a pointed $(F \times \Delta_X \times \Delta_Q)$-coalgebra $\mathsf{R} := (R, \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ with $R := X \times Q$, $\rho_X := \mathsf{pr}_1$ and $\rho_Q := \mathsf{pr}_2$. The marking μ is used to define $r_I := (x_I, \mathsf{in}_1 \circ \mu(x_I))$ and:

$$\rho_F := X \times Q \xrightarrow{\mathsf{pr}_1} X \xrightarrow{\xi} FX \xrightarrow{\boxdot_X} F(F'X \times X) \xrightarrow{F(F'\mu \times \langle id, \mu \rangle)} F(F'C \times (X \times C))$$
$$\xrightarrow{\cong} F(X \times (F'C \times C)) \xrightarrow{F(id \times \mathsf{in}_2)} F(X \times Q).$$

It can be verified that the reachable subcoalgebra of R is a pre-run, using Lemma 2.5 and properties of the marking μ. $\square$

The benefit of working with markings instead of (pre-)runs is that markings are defined solely in terms of the algebra, as opposed to in terms of the algebraic automaton. We will see in Lemma 5.8 that by equipping (X, ξ) with a suitable $(F + G)$-coalgebra structure, markings turn into $(F + G)$-coalgebra-to-algebra morphisms. This will allow us to find existence and uniqueness properties of markings that follow from the recursive structure of thin behaviours. Consequently, Proposition 5.3 will allow us to draw conclusions about pre-runs of the algebraic automaton.

Properties of Markings

The first property of markings is that every thin coalgebra can be marked. Concretely, for all thin coalgebras (X, ξ) and all coherent algebras (C, γ), we show that $\mu := cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$ is a marking of (X, ξ) with (C, γ). Our strategy is to show that $cev_{(C,\gamma)}$ is a marking and that markings are preserved under precomposition with F-coalgebra morphisms.

Lemma 5.4 *If (C, γ) is a coherent $(F + G)$-algebra, then $cev_{(C,\gamma)} : (Z^\flat, \beta^\flat) \to (C, \gamma)$ is a marking of $(Z^\flat, \zeta^\flat)$ with (C, γ).*

Proof. To see that $cev_{(C,\gamma)}$ satisfies condition (i) of markings, i.e., $cev_{(C,\gamma)} : (Z^\flat, \zeta^\flat) \to (C, \gamma_0)$ is an F-coalgebra-to-algebra morphism, consider the diagram to the right. We have:
$$cev_{(C,\gamma)} \circ \beta_0^\flat = \gamma_0 \circ F cev_{(C,\gamma)} = \gamma_0 \circ F cev_{(C,\gamma)} \circ \zeta^\flat \circ \beta_0^\flat,$$
where the first equality uses that $cev_{(C,\gamma)}$ is an $(F+G)$-algebra morphism and the second equality uses Equation (1). Now since $\beta_0^\flat$ is epic, we conclude $cev_{(C,\gamma)} = \gamma_0 \circ F cev_{(C,\gamma)} \circ \zeta^\flat$, i.e., $cev_{(C,\gamma)}$ is an F-coalgebra-to-algebra morphism.

To see that $cev_{(C,\gamma)}$ satisfies condition (ii) of markings, let $(z_n)_{n \in \omega} \in (Z^\flat)^\omega$ and $(\bar{z}'_n)_{n>0} \in GZ^\flat$ satisfy $\triangleright_{Z^\flat}(\bar{z}'_{n+1}, z_{n+1}) = \zeta^\flat(z_n)$, for all $n \in \omega$. It follows from Equation (2) that $\beta_1^\flat((\bar{z}'_n)_{n>0}) = z_0$. Hence:

$$cev_{(C,\gamma)}(z_0) = cev_{(C,\gamma)}(\beta_1^\flat((\bar{z}'_n)_{n>0})) = (\gamma_1 \circ G cev_{(C,\gamma)})((\bar{z}'_n)_{n>0}). \qquad \square$$

Lemma 5.5 *If $\mu : (X, \xi) \to (C, \gamma)$ is a marking and $f : (Y, \upsilon) \to (X, \xi)$ is an F-coalgebra morphism, then $\mu \circ f$ is a marking.*

12

Proposition 5.6 (Existence of Markings) *For every thin F-coalgebra (X, ξ) and every coherent $(F + G)$-algebra (C, γ), there exists a marking μ of (X, ξ) with (C, γ) given by $\mu = cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$.*

$$
\begin{array}{ccccc}
X & \xrightarrow{tbeh_{(X,\xi)}} & Z^\flat & \xrightarrow{cev_{(C,\gamma)}} & C \\
{\scriptstyle \xi}\big\downarrow & & {\scriptstyle \zeta^\flat}\big\uparrow & & {\scriptstyle \gamma_0}\big\uparrow \\
FX & \xrightarrow[F\,tbeh_{(X,\xi)}]{} & FZ^\flat & \xrightarrow[F\,cev_{(C,\gamma)}]{} & FC
\end{array}
$$

Proof. By Lemma 5.4, $cev_{(C,\gamma)} : Z^\flat \to C$ is a marking of $(Z^\flat, \zeta^\flat)$ with (C, γ). By Lemma 5.5, $cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$ is a marking of (X, ξ) with (C, γ). □

The second central property of markings is uniqueness: there do not exist two distinct markings of a given thin coalgebra with a given coherent algebra (Proposition 5.9). The key insight behind the proof is that every thin coalgebra (X, ξ) can be transformed into a recursive $(F + G)$-coalgebra such that markings of (X, ξ) become $(F + G)$-coalgebra-to-algebra morphisms (Lemma 5.8). Uniqueness of markings will then follow from the fact that coalgebra-to-algebra morphisms with a recursive domain coincide.

The recursive $(F+G)$-coalgebra structure on X is inherited from a canonical recursive $(F+G)$-coalgebra structure η on $Z^\flat$. Intuitively, η decomposes a normal term $z \in Z^\flat$ into its normal subterms.

Definition 5.7 Let $\iota : Z^\flat \to A$ be the map sending each thin behaviour to its unique normal representative. Define an $(F + G)$-coalgebra structure η on $Z^\flat$ by $\eta := (F + G)[\![-]\!] \circ \alpha^{-1} \circ \iota$.

$$
\begin{array}{ccc}
A & \underset{[\![-]\!]}{\overset{\iota}{\rightleftarrows}} & Z^\flat \\
{\scriptstyle \alpha^{-1}}\big\downarrow & & \big\downarrow{\scriptstyle \eta} \\
(F + G)A & \xrightarrow[(F+G)[\![-]\!]]{} & (F + G)Z^\flat
\end{array}
$$

Next, we show that we can define an $(F + G)$-coalgebra structure on any thin coalgebra that turns markings into $(F + G)$-coalgebra-to-algebra morphisms.

Lemma 5.8 *Let (X, ξ) be a thin F-coalgebra and (C, γ) be a coherent $(F + G)$-algebra. There exists an $(F + G)$-coalgebra structure υ on X such that:*

(i) $tbeh := tbeh_{(X,\xi)}$ is an $(F + G)$-coalgebra morphism from (X, υ) to $(Z^\flat, \eta)$, and

(ii) every marking $\mu : X \to C$ is an $(F + G)$-coalgebra-to-algebra morphism $(X, \upsilon) \to (C, \gamma)$.

$$
\begin{array}{ccccc}
 & & FX & \xrightarrow{F\,tbeh} & FZ^\flat \\
 & & {\scriptstyle \xi}\big\uparrow & & {\scriptstyle \zeta^\flat}\big\uparrow \\
C & \xleftarrow{\mu} & X & \xrightarrow{tbeh} & Z^\flat \\
{\scriptstyle \gamma}\big\uparrow & & {\scriptstyle \upsilon}\big\downarrow & & \big\downarrow{\scriptstyle \eta} \\
(F + G)C & \xleftarrow[(F+G)\mu]{} & (F + G)X & \xrightarrow[(F+G)\,tbeh]{} & (F + G)Z^\flat
\end{array}
$$

Proposition 5.9 (Uniqueness of Markings) *For every thin F-coalgebra (X, ξ) and every coherent $(F + G)$-algebra (C, γ), there is at most one marking of (X, ξ) with (C, γ).*

Proof. Let μ_1 and μ_2 be two markings of (X, ξ) with (C, γ). By appealing to Lemma 5.8, we obtain a

coalgebra structure $\upsilon : X \to (F + G)X$. Consider the diagram:

$$
\begin{array}{ccccccc}
C & \xleftarrow{\;\;\mu_1\;\;\atop\mu_2\;} & X & \xrightarrow{\;tbeh_{(X,\xi)}\;} & Z^\flat & \xrightarrow{\;\;\iota\;\;} & A \\
\gamma\uparrow & & \downarrow\upsilon & & \downarrow\eta & & \downarrow\alpha^{-1} \\
(F+G)C & \xleftarrow[\;(F+G)\mu_2\;]{\;(F+G)\mu_1\;} & (F+G)X & \xrightarrow[(F+G)tbeh_{(X,\xi)}]{} & (F+G)Z^\flat & \xrightarrow[(F+G)\iota]{} & (F+G)A
\end{array}
$$

We know μ_1, μ_2 are $(F + G)$-coalgebra-to-algebra morphisms and that $tbeh_{(X,\xi)}$, ι are $(F + G)$-coalgebra morphisms. Observe that, since (A, α) is an initial $(F + G)$-algebra, the coalgebra (A, α^{-1}) is recursive [2, Corollary 8.2]. Moreover, any coalgebra mapping into a recursive coalgebra is also recursive [2, Corollary 8.2], hence (X, υ) is recursive. Now μ_1 and μ_2 are two coalgebra-to-algebra morphisms with a recursive coalgebra as their domain, therefore $\mu_1 = \mu_2$. □

In the proof of Proposition 5.9, note the instrumental role of the inductive structure of thin behaviours. It is what allowed us to obtain a recursive $(F + G)$-coalgebra structure on X.

5.3 Acceptance and Unambiguity of the Algebraic Automaton

Theorem 5.10 *Let $\mathsf{C} = (C, \gamma = [\gamma_0, \gamma_1], U)$ be a coherent $(F+G)$-algebra with a recognising set. For every thin pointed F-coalgebra (X, ξ, x_I), the behaviour of x_I is in the language of C if and only if the algebraic automaton $\mathcal{A}_\mathsf{C}$ accepts (X, ξ, x_I).*

Proof. Suppose $(cev_{(C,\gamma)} \circ tbeh_{(X,\xi)})(x_I) \in U$. By Proposition 5.6, $\mu := cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$ is a marking of (X, ξ) with (C, γ). By Proposition 5.3 (ii), there exists a pre-run $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I) with $\rho_Q(r_I) \in \mathsf{in}_1[C]$ and $[id, \mathsf{pr}_2] \circ \rho_Q = \mu \circ \rho_X$.

Hence $[id, \mathsf{pr}_2] \circ \rho_Q = \mu \circ \rho_X = cev_{(C,\gamma)} \circ tbeh_{(X,\xi)} \circ \rho_X$. We conclude that $\rho_Q(r_I) \in (\mathsf{in}_1 \circ cev_{(C,\gamma)} \circ tbeh_{(X,\xi)})(x_I) \in \mathsf{in}_1[U]$, i.e., (R, ρ, r_I) is an accepting pre-run. By Lemma 3.7, (X, ξ, x_I) is accepted by $\mathcal{A}_\mathsf{C}$.

$$
\begin{array}{ccccc}
R & \xrightarrow{\;\rho_X\;} & X & \xrightarrow{\;tbeh_{(X,\xi)}\;} & Z^\flat \\
\rho_Q\downarrow & & & \searrow^{\mu} & \downarrow^{cev_{(C,\gamma)}} \\
Q & & \xrightarrow[{[id,\mathsf{pr}_2]}]{} & & C
\end{array}
$$

Conversely, suppose there exists an accepting run $(R, \rho = \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I). By Proposition 5.3 (i), the map $\mu = [id, \mathsf{pr}_2] \circ \rho_Q$ is a marking of (R, ρ_F) with (C, γ). By Propositions 5.6 and 5.9, we have $\mu = cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$. Hence:

$$
\begin{array}{ccc}
R & \xrightarrow{\;\;\mu\;\;} & C \\
\rho_X\downarrow & \searrow^{tbeh_{(R,\rho_F)}} & \uparrow^{cev_{(C,\gamma)}} \\
X & \xrightarrow[tbeh_{(X,\xi)}]{} & Z^\flat
\end{array}
$$

$$
(cev_{(C,\gamma)} \circ tbeh_{(X,\xi)})(x_I) = (cev_{(C,\gamma)} \circ tbeh_{(X,\xi)} \circ \rho_X)(r_I) = \\
(cev_{(C,\gamma)} \circ tbeh_{(R,\rho_F)})(r_I) = \mu(r_I) = ([id, \mathsf{pr}_2] \circ \rho_Q)(r_I).
$$

But since (R, ρ, r_I) is accepting, $\rho_Q(r_I) \in \mathsf{in}_1(U)$, so $(cev_{(C,\gamma)} \circ tbeh_{(X,\xi)})(x_I) = ([id, \mathsf{pr}_2] \circ \rho_Q)(r_I) \in U$. □

In order to prove unambiguity of the algebraic automaton, we first show that in every pre-run ρ_Q is uniquely determined by ρ_F and ρ_X.
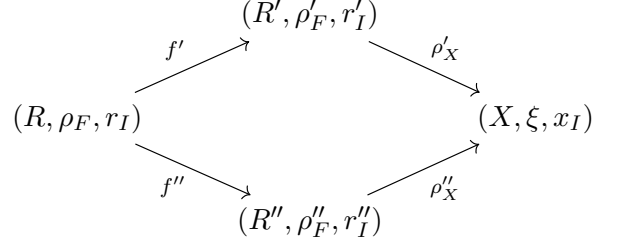
Lemma 5.11 *Let C be a coherent $(F + G)$-algebra with a recognising set, and (X, ξ, x_I) be a thin pointed F-coalgebra. If $(R, \langle \rho_F, \rho_X, \rho_Q \rangle, r_I)$ and $(R, \rho_F, \rho_X, \rho_Q' \rangle, r_I)$ are pre-runs of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I), then $\rho_Q = \rho_Q'$.*

Proof (Sketch) Suppose $\mathsf{C} = (C, \gamma, U)$. By Proposition 5.3 (i), $[id, \mathsf{pr}_2] \circ \rho_Q : R \to C$ is a marking of (R, ρ_F) with (C, γ). By Propositions 5.6 and 5.9, $[id, \mathsf{pr}_2] \circ \rho_Q = cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$. Similarly, $[id, \mathsf{pr}_2] \circ \rho_Q' = cev_{(C,\gamma)} \circ tbeh_{(X,\xi)}$, so $[id, \mathsf{pr}_2] \circ \rho_Q = [id, \mathsf{pr}_2] \circ \rho_Q'$. Now it can be shown by induction on the successor relation of (R, ρ_F, r_I) that $\rho_Q = \rho_Q'$. □

Theorem 5.12 *Let C be a coherent $(F + G)$-algebra with a recognising set, and (X, ξ, x_I) be a thin pointed F-coalgebra. The algebraic automaton $\mathcal{A}_\mathsf{C}$ is unambiguous on (X, ξ, x_I).*

Proof. Let $(R', \rho' = \langle \rho_F', \rho_X', \rho_Q' \rangle, r_I')$ and $(R'', \rho'' = \langle \rho_F'', \rho_X'', \rho_Q'' \rangle, r_I'')$ be two accepting runs of $\mathcal{A}_\mathsf{C}$ on (X, ξ, x_I). We prove equality between these runs by exhibiting a span between them.

Since $\rho_X' : (R', \rho_F', r_I') \to (X, \xi, x_I)$ and $\rho_X'' : (R'', \rho_F'', r_I'') \to (X, \xi, x_I)$ are pointed F-coalgebra morphisms and F preserves weak pullbacks, there exists a reachable pointed F-coalgebra (R, ρ_F, r_I) with F-coalgebra morphisms $f' : (R, \rho_F, r_I) \to (R', \rho_F', r_I')$ and $f'' : (R, \rho_F, r_I) \to (R'', \rho_F'', r_I'')$ such that $\rho_X' \circ f' = \rho_X'' \circ f''$.

$$
\begin{array}{ccc}
 & (R', \rho_F', r_I') & \\
\nearrow f' & & \searrow \rho_X' \\
(R, \rho_F, r_I) & & (X, \xi, x_I) \\
\searrow f'' & & \nearrow \rho_X'' \\
 & (R'', \rho_F'', r_I'') &
\end{array}
$$

Define $\rho_X := \rho_X' \circ f'$. It can be shown that $(R, \langle \rho_F, \rho_X, \rho_Q' \circ f' \rangle, r_I)$ and $(R, \langle \rho_F, \rho_X, \rho_Q'' \circ f'' \rangle, r_I)$ are pre-runs, so according to Lemma 5.11, $\rho_Q' \circ f' = \rho_Q'' \circ f''$. Hence, by setting $\rho_Q := \rho_Q' \circ f'$, we have that $f' : (R, \langle \rho_F, \rho_X, \rho_Q \rangle, r_I) \to (R', \rho', r_I')$ and $f'' : (R, \langle \rho_F, \rho_X, \rho_Q \rangle, r_I) \to (R'', \rho'', r_I'')$ are $(F \times \Delta_X \times \Delta_Q)$-morphisms. Consequently, $((R, \rho, r_I), f', f'')$ is a span in the category of pointed $(F \times \Delta_X \times \Delta_Q)$-coalgebras. This means that (R', ρ', r_I') and (R'', ρ'', r_I'') are behaviourally equivalent pointed subcoalgebras of the final coalgebra, hence they are equal. $\qquad\square$

6 Combining the Two Constructions

Here we derive the main results of the paper, by employing the automaton algebra and the algebraic automaton constructions. We begin by showing that, when restricted to thin coalgebras, every parity F-automaton has an equivalent unambiguous parity F-automaton. We will make use of a property that we hinted at earlier: that rational algebras induce automata with ω-regular acceptance.

Lemma 6.1 *Let $\mathsf{C} = (C, \gamma, U)$ be a rational $(F + G)$-algebra with a recognising set. Then the acceptance condition of the algebraic automaton $\mathcal{A}_\mathsf{C}$ is ω-regular.*

Proof (Sketch) It follows from the ω-semigroup structure on (C, γ) that, for every $c \in C$, the language $L(c) := \{(\bar{c}_n')_{n \in \omega} \in (F'C)^\omega \mid \gamma_1((\bar{c}_n')_{n \in \omega}) = c\}$ is ω-regular. This can be used to show that the acceptance condition $\mathrm{Acc} = \{\mathsf{in}_1(c_0) \cdot (\mathsf{in}_2(\bar{c}_n', c_n))_{n>0} \mid \forall m (c_m = \gamma_1((\bar{c}_n')_{n>m}))\}$ is ω-regular. $\qquad\square$

Theorem 6.2 *For every parity F-automaton $\mathcal{A}$, there exists a parity automaton $\mathcal{A}'$ such that:*

(i) *$\mathcal{A}$ and $\mathcal{A}'$ accept the same thin F-coalgebras, and*

(ii) *$\mathcal{A}'$ is unambiguous on thin F-coalgebras.*

Proof. Since parity conditions are prefix-agnostic, by Theorem 4.3, the automaton algebra $\mathsf{C}_\mathcal{A}$ accepts exactly those thin behaviours accepted by $\mathcal{A}$. By Proposition 4.6, $\mathsf{C}_\mathcal{A}$ is rational, so, by Lemma 6.1, its algebraic automaton $\mathcal{B} := \mathcal{A}_{\mathsf{C}_\mathcal{A}}$ has an ω-regular acceptance condition. By Theorem 5.10, $\mathcal{B}$ accepts the same thin coalgebras as $\mathcal{A}$, while by Theorem 5.12, $\mathcal{B}$ is unambiguous on thin coalgebras. Finally, applying Proposition 3.9 to $\mathcal{B}$ gives us the desired automaton $\mathcal{A}'$. $\qquad\square$

As our second main result, we give an automaton-theoretic characterisation of languages of finite coherent $(F + G)$-algebras. Concretely, we show that coherent algebras are as expressive as automata with prefix-agnostic acceptance (restricted to thin coalgebras). The key observation is that the acceptance condition of the algebraic automaton can be adjusted to a prefix-agnostic condition.

Lemma 6.3 *Let $\mathsf{C} = (C, \gamma, U)$ be a finite coherent $(F + G)$-algebra with a recognising set. There exists an F-automaton with a prefix-agnostic acceptance condition whose runs coincide with the runs of $\mathcal{A}_\mathsf{C}$.*

Proof (Sketch) Suppose $\mathcal{A}_\mathsf{C} = (Q, \delta, Q_I, \mathrm{Acc})$. Using coherence of C, the following automaton can be shown to satisfy the desired conditions: $\mathcal{A}' := (Q, \delta, Q_I, \mathrm{Acc}')$ with:

$$
\mathrm{Acc}' := \{(q_n)_{n \in \omega} \in Q^\omega \mid \exists m \in \omega, (\bar{c}_n')_{n \geq m} \in F'C, (c_n)_{n \geq m} \in C^\omega :
$$
$$
(q_n)_{n \geq m} = (\mathsf{in}_2(\bar{c}_n', c_n))_{n \geq m} \wedge \forall k \geq m (c_m = \gamma_1((\bar{c}_n')_{n \geq k}))\}. \quad\square
$$

Theorem 6.4 *Restricted to thin F-coalgebras, finite coherent $(F + G)$-algebras recognise exactly the languages accepted by F-automata with a prefix-agnostic acceptance condition. More precisely, a language L of thin F-behaviours equals $L(\mathsf{C})$, for some finite coherent $(F + G)$-algebra C, if and only if L consists of those thin F-behaviours accepted by $\mathcal{A}$, for some F-automaton $\mathcal{A}$ with prefix-agnostic acceptance.*

Proof. Given a finite coherent $(F+G)$-algebra C with a recognising set, we have by Theorem 5.10 that the language of C consists precisely of the thin coalgebras accepted by its algebraic automaton. By Lemma 6.3, there exists an equivalent prefix-agnostic automaton. Conversely, for every F-automaton with prefix-agnostic acceptance, by Theorem 4.3, its automaton algebra recognises precisely those thin coalgebras accepted by the automaton. □

7 Conclusion

In this paper, we saw how to connect F-automata with prefix-agnostic acceptance to finite coherent $(F + G)$-algebras in order to transform an arbitrary F-automaton into an unambiguous one. We gave two constructions: the automaton algebra and the algebraic automaton constructions, both of which generalise the corresponding classical constructions for thin trees [16]. In order to prove unambiguity of the algebraic automaton (Theorem 5.12), we linked algebraic automaton pre-runs to markings, which are F-coalgebra-to-algebra morphisms with an extra condition. We used the inductive structure of thin behaviours [5] to show existence and uniqueness of markings, which implied existence and uniqueness of runs. We concluded from the two constructions that finite $(F + G)$-algebras recognise the same languages of thin F-behaviours as F-automata with prefix-agnostic acceptance (Theorem 6.4).

In applications, one usually considers parity F-automata. Hence we showed that the unambiguous automaton obtained from a parity automaton is itself a parity automaton (Theorem 6.2). To this end, we identified rational $(F + G)$-algebras as a subclass of coherent algebras that correspond to parity automata. We observed that rational algebras generalise thin algebras [16] to analytic functors.

In addition to providing a useful generalisation (beyond trees and polynomial functors) of an existing construction, our use of the context decomposition operator in the definition of the algebraic automaton, and our key insight that markings (as defined in [16]) correspond to $(F+G)$-coalgebra-to-algebra morphisms shed new light on the original construction in loc. cit. and on the reasons it delivers unambiguity.

A natural direction for future work is to incorporate our unambiguous parity automaton construction into model-checking algorithms, such as the one proposed in [6]. In this context, the size of the resulting automaton is crucial. In principle, our constructions yield at least an exponential blow-up but simple optimisations such as removing unreachable states could considerably improve the automaton size.

We defined recognition by coherent algebra only for languages of thin F-behaviours, but some of our constructions can be extended to all F-behaviours. In particular, by considering $(F+G)$-algebra morphisms from (Z, β) to a finite coherent $(F+G)$-algebra, we obtain a notion of recognition for arbitrary F-behaviours. We can extend Theorem 4.3 to show that the automaton algebra recognises the same language over all F-behaviours. In contrast, the properties of the algebraic automaton make essential use of thinness and, without it, neither Theorem 5.10 nor Theorem 5.12 seem to hold. We leave further investigations into coherent algebra recognition of non-thin behaviours as future work.

Finally, our characterisation of languages of finite coherent $(F+G)$-algebras, together with Example 3.3, show that the expressivity of these algebras lies beyond the tractable realm of regular languages. We "corrected" this by equipping the algebras with additional structure, thus obtaining rational algebras. Yet, rationality only played a role in ensuring that the algebraic automaton has a parity condition. This suggests that coherent algebras could also be specialised with alternative additional structure in order to study different classes of languages, while maintaining the correctness of the unambiguity construction.

References

[1] Abbott, M., T. Altenkirch, C. McBride and N. Ghani, *∂ for data: Differentiating data structures*, Fundamenta Informaticae **65**, pages 1–28 (2005).
https://dl.acm.org/doi/10.5555/2370077.2370079

[2] Adámek, J., S. Milius and L. Moss, *On well-founded and recursive coalgebras*, in: J. Goubault-Larrecq and B. König, editors, *Foundations of Software Science and Computation Structures*, pages 17–36, Springer International Publishing (2020).

[3] Baier, C. and J.-P. Katoen, *Principles of model checking*, Comput. J. **53**, page 615–616 (2010).
https://doi.org/10.1093/comjnl/bxp025

[4] Carayol, A., C. Löding, D. Niwiński and I. Walukiewicz, *Choice functions and well-orderings over the infinite binary tree*, Central European Journal of Mathematics **8**, pages 662—-682 (2010).

[5] Chernev, A., C. Cîrstea, H. H. Hansen and C. Kupke, *Thin coalgebraic behaviours are inductive.* To appear in LICS 2025. Preliminary version available on arXiv.
https://doi.org/10.48550/arXiv.2504.07013

[6] Cîrstea, C. and C. Kupke, *Measure-theoretic semantics for quantitative parity automata*, in: B. Klin and E. Pimentel, editors, *31st EACSL Annual Conference on Computer Science Logic (CSL 2023)*, volume 252 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:20, Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2023).
https://doi.org/10.4230/LIPICS.CSL.2023.14

[7] Cirstea, C., S. Shimizu and I. Hasuo, *Parity Automata for Quantitative Linear Time Logics*, in: F. Bonchi and B. König, editors, *7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017)*, volume 72 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:18, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2017), ISBN 978-3-95977-033-0, ISSN 1868-8969.
https://doi.org/10.4230/LIPIcs.CALCO.2017.7

[8] Grädel, E., W. Thomas and T. Wilke, editors, *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *Lecture Notes in Computer Science*, Springer (2002).
https://doi.org/10.1007/3-540-36387-4

[9] Gumm, H., *From t-coalgebras to filter structures and transition systems*, in: J. Fiadeiro, N. Harman, M. Roggenbach and J. Rutten, editors, *Algebra and Coalgebra in Computer Science*, pages 194–212, Springer Berlin Heidelberg (2005).

[10] Hasegawa, R., *Two applications of analytic functors*, Theoretical Computer Science **272**, pages 113–175 (2002).
https://doi.org/10.1016/S0304-3975(00)00349-2

[11] Joyal, A., *Foncteurs analytiques et espèces de structures*, in: G. Labelle and P. Leroux, editors, *Combinatoire énumérative*, pages 126–159, Springer Berlin Heidelberg (1986).
https://doi.org/10.1007/BFb0072514

[12] Kupke, C. and Y. Venema, *Coalgebraic automata theory: Basic results*, Logical Methods in Computer Science **4** (2008).
https://doi.org/10.2168/LMCS-4(4:10)2008

[13] Kurz, A. and J. Velebil, *Relation lifting, a survey*, Journal of Logical and Algebraic Methods in Programming **85**, pages 475–499 (2016). Relational and algebraic methods in computer science.
https://doi.org/https://doi.org/10.1016/j.jlamp.2015.08.002

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

[15] Rutten, J., *Universal coalgebra: A theory of systems*, Theoretical Computer Science **249**, pages 3–80 (2000).

[16] Skrzypczak, M., *Recognition by Thin Algebras*, pages 121–135, Springer Berlin Heidelberg (2016).
https://doi.org/10.1007/978-3-662-52947-8_7

[17] Urabe, N., S. Shimizu and I. Hasuo, *Coalgebraic Trace Semantics for Buechi and Parity Automata*, in: J. Desharnais and R. Jagadeesan, editors, *27th International Conference on Concurrency Theory (CONCUR 2016)*, volume 59 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 24:1–24:15, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2016), ISBN 978-3-95977-017-0, ISSN 1868-8969.
https://doi.org/10.4230/LIPIcs.CONCUR.2016.24