Arboreal Categories and Resources

Samson Abramsky
Department of Computer Science, University of Oxford, UK

Luca Reggio
Department of Computer Science, University of Oxford, UK

Abstract

We introduce *arboreal categories*, which have an intrinsic process structure, allowing dynamic notions such as bisimulation and back-and-forth games, and resource notions such as number of rounds of a game, to be defined. These are related to extensional or "static" structures via *arboreal covers*, which are resource-indexed comonadic adjunctions. These ideas are developed in a very general, axiomatic setting, and applied to relational structures, where the comonadic constructions for pebbling, Ehrenfeucht-Fraïssé and modal bisimulation games recently introduced in [1, 5, 6] are recovered, showing that many of the fundamental notions of finite model theory and descriptive complexity arise from instances of arboreal covers.

2012 ACM Subject Classification Theory of computation → Categorical semantics; Theory of computation → Finite Model Theory

Keywords and phrases factorisation system, embedding, comonad, coalgebra, open maps, bisimulation, game, resources, relational structures, finite model theory

Digital Object Identifier 10.4230/LIPIcs.ICALP.2021.115

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Related Version *Extended Version*: `https://arxiv.org/abs/2102.08109`

Funding *Samson Abramsky*: Research supported by the EPSRC project EP/T00696X/1 "Resources and Co-Resources: a junction between categorical semantics and descriptive complexity".
Luca Reggio: Research supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 837724.

Acknowledgements Feedback from Tomáš Jakl and Dan Marsden is gratefully acknowledged.

1 Introduction

In previous work ([1, 5, 6]), it has been shown how a range of model comparison games which play a central role in finite model theory, including Ehrenfeucht-Fraïssé, pebbling, and bisimulation games, can be captured in terms of resource-indexed comonads on the category of relational structures and homomorphisms. This was done for k-pebble games in [1], and extended to Ehrenfeucht-Fraïssé games, and bisimulation games for the modal fragment, in [5]. In subsequent work, this has been further extended to games for generalized quantifiers [9], and for guarded fragments of first-order logic [3]. An important feature of this comonadic analysis is that it leads to novel characterisations of important combinatorial parameters such as tree-width and tree-depth. The coalgebras for each of these comonads correspond to certain forms of tree decompositions of structures, with the resource index matching the corresponding combinatorial parameter.

This leads to the question motivating the present paper:

$\boxed{\text{Can we capture the significant common elements of these constructions?}}$

48th International Colloquium on Automata, Languages, and Programming (ICALP 2021).
Editors: Nikhil Bansal, Emanuela Merelli, and James Worrell; Article No. 115; pp. 115:1–115:20
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Our aim is to develop an elegant axiomatic account, based on clear conceptual principles, which will yield all these examples and more, and allow a deeper and more general understanding of resources.

Conceptually, a key ingredient is the assignment of a process structure – an intensional description – to an extensional object, such as a function, a set, or a relational structure. It is this process structure, unfolding in space and time, to which a resource parameter can be applied, which can then be transferred to the extensional object. At the basic level of computability, this happens when we assign a Turing machine description or a Gödel number to a recursive function. It is then meaningful to assign a complexity measure to the function. The same phenomenon arises in semantics: for example, the notion of *sequentiality* is applicable to a *process* computing a higher-order function. Reifying these processes in the form of *game semantics* led to a resolution of the famous full abstraction problem for PCF [2, 11], and to a wealth of subsequent results [20].

It is now becoming clear that this phenomenon is at play in the game comonads described in [1, 5, 6, 9, 3]. They build tree-structured covers of a given, purely extensional relational structure. Such a tree cover will in general not have the full properties of the original structure, but be a "best approximation" in some resource-restricted setting. More precisely, this means that we have an adjunction, yielding the corresponding comonad. The objects of the category where the approximations live have an intrinsic tree structure, which can be captured axiomatically. The tree encodes a process for generating (parts of) the relational structure, to which resource notions can be applied.

In this paper, we make this intuition precise. We introduce a notion of *arboreal category*, and show how all the examples of game comonads considered to date arise from *arboreal covers*, i.e. adjunctions between extensional categories of relational structures, and arboreal categories. Importantly, these adjunctions are comonadic, and the categories of coalgebras provide a setting for a general notion of bisimulation, which yields a wide range of logical equivalences in the examples. This notion refines the open maps formulation of bisimulation [13, 12] with the condition that the maps are *pathwise embeddings*, generalizing the ideas introduced in [6]. This allows a much wider range of logical equivalences to be captured.

After some preliminaries, we shall develop the axiomatization of paths, open pathwise embeddings and bisimulations, and arboreal categories. Then we establish the correspondence between bisimulations and back-and-forth equivalences in the setting of arboreal categories. Next, we show how many of the fundamental notions of finite model theory and descriptive complexity arise from instances of arboreal covers. We shall use the concrete constructions in finite model theory as running examples throughout.

We conclude this introduction by observing that the notion of *extendability*, a key ingredient for Rossman-type preservation theorems [19], can be defined in this general setting (more details are provided in the expanded version [4]).

2 Preliminaries

We shall assume familiarity with standard notions in category theory. All needed background can be found in [7, 16]. All categories under consideration are assumed to be locally small and *well-powered*, i.e. every object has a set of subobjects (as opposed to a proper class).

▶ **Example 1.** The extensional categories of primary interest in this paper are categories of relational structures. A relational vocabulary σ is a set of relation symbols R, each with a specified positive integer arity. A σ-structure $\mathcal{A}$ is given by a set A, the universe of the structure, and for each R in σ with arity n, a relation $R^{\mathcal{A}} \subseteq A^n$. A homomorphism

$h\colon \mathcal{A} \to \mathcal{B}$ is a function $h\colon A \to B$ such that, for each relation symbol R of arity n in σ, for all $a_1, \ldots, a_n$ in A, $R^{\mathcal{A}}(a_1, \ldots, a_n) \Rightarrow R^{\mathcal{B}}(h(a_1), \ldots, h(a_n))$. We write $\mathsf{Struct}(\sigma)$ for the category of σ-structures and homomorphisms. The Gaifman graph of a structure $\mathcal{A}$ is a graph with vertices A, such that two distinct elements are adjacent if they both occur in some tuple $\vec{a} \in R^{\mathcal{A}}$ for some relation symbol R in σ.

2.1 Proper factorisation systems

We recall the notion of weak factorisation system in a category $\mathcal{C}$. Given arrows e and m in $\mathcal{C}$, we say that e has the *left lifting property* with respect to m, or that m has the *right lifting property* with respect to e, if for every commutative square as on the left-hand side below

$$\begin{tikzcd} \bullet \arrow[r, "e"] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r, "m"] & \bullet \end{tikzcd} \qquad\qquad \begin{tikzcd} \bullet \arrow[r, "e"] \arrow[d] & \bullet \arrow[d] \arrow[dl, "d"'] \\ \bullet \arrow[r, "m"] & \bullet \end{tikzcd}$$

there exists a (not necessarily unique) *diagonal filler*, i.e. an arrow d such that the right-hand diagram above commutes. If this is the case, we write $e \pitchfork m$. For any class $\mathcal{H}$ of morphisms in $\mathcal{C}$, let ${}^{\pitchfork}\mathcal{H}$ (respectively $\mathcal{H}^{\pitchfork}$) be the class of morphisms having the left (respectively right) lifting property with respect to every morphism in $\mathcal{H}$.

▶ **Definition 2.** *A pair of classes of morphisms* $(\mathcal{Q}, \mathcal{M})$ *in a category* $\mathcal{C}$ *is a* weak factorisation system *provided it satisfies the following conditions:*

- **(i)** *every morphism* f *in* $\mathcal{C}$ *can be written as* $f = m \circ e$ *with* $e \in \mathcal{Q}$ *and* $m \in \mathcal{M}$*;*
- **(ii)** $\mathcal{Q} = {}^{\pitchfork}\mathcal{M}$ *and* $\mathcal{M} = \mathcal{Q}^{\pitchfork}$.

A proper factorisation system *is a weak factorisation system* $(\mathcal{Q}, \mathcal{M})$ *such that* $\mathcal{Q} \subseteq \{\text{epis}\}$ *and* $\mathcal{M} \subseteq \{\text{monos}\}$*. A proper factorisation system is* stable *if, for any* $e \in \mathcal{Q}$ *and* $m \in \mathcal{M}$ *with common codomain, the pullback of* e *along* m *exists and belongs to* $\mathcal{Q}$*.*[1]

▶ Remark 3. Any proper factorisation system is an *orthogonal* factorisation system, i.e. the diagonal fillers are unique. In particular, factorisations are unique up to (unique) isomorphism.

▶ **Example 4.** If $\mathcal{A}$ is a relational structure, then for any $S \subseteq A$, there is an induced substructure with universe S. The inclusion map $S \hookrightarrow A$ is an *embedding*, i.e. an injective homomorphism which reflects as well as preserves relations. Any embedding $m\colon \mathcal{A} \to \mathcal{B}$ factors as $\mathcal{A} \cong \mathsf{Im}(m) \hookrightarrow \mathcal{B}$. Taking $\mathcal{Q}$ to be the surjective homomorphisms and $\mathcal{M}$ to be the embeddings gives a proper factorisation system on $\mathsf{Struct}(\sigma)$. This factorisation system is stable because pullbacks in $\mathsf{Struct}(\sigma)$ are computed in the category of sets and functions, where (surjections, injections) is a stable proper factorisation system.

Next, we state some well known properties of weak factorisation systems (cf. [10] or [18]):

▶ **Lemma 5.** *Let* $(\mathcal{Q}, \mathcal{M})$ *be a weak factorisation system in* $\mathcal{C}$*. The following hold:*

- **(a)** $\mathcal{Q}$ *and* $\mathcal{M}$ *are closed under compositions.*
- **(b)** $\mathcal{Q} \cap \mathcal{M} = \{$*isomorphisms*$\}$.
- **(c)** *The pullback in* $\mathcal{C}$ *of an* $\mathcal{M}$*-morphism along any morphism, if it exists, is again in* $\mathcal{M}$*.*

Moreover, if $(\mathcal{Q}, \mathcal{M})$ *is proper, then the following hold:*

- **(d)** $g \circ f \in \mathcal{Q}$ *implies* $g \in \mathcal{Q}$.
- **(e)** $g \circ f \in \mathcal{M}$ *implies* $f \in \mathcal{M}$.

[1] In the literature, the adjective *stable* is usually reserved for the stronger property stating that, for every $e \in \mathcal{Q}$, the pullback of e along any morphism exists and belongs to $\mathcal{Q}$.

Throughout this paper, we will refer to $\mathcal{M}$-morphisms as *embeddings* and denote them by $\rightarrowtail$. $\mathcal{Q}$-morphisms will be referred to as *quotients* and denoted by $\twoheadrightarrow$.

Assume $\mathcal{C}$ is a category admitting a proper factorisation system $(\mathcal{Q}, \mathcal{M})$. In the same way that one usually defines the poset of subobjects of a given object $X \in \mathcal{C}$, we can define the poset of $\mathcal{M}$-subobjects of X. Given embeddings $m\colon S \rightarrowtail X$ and $n\colon T \rightarrowtail X$, let us say that $m \trianglelefteq n$ provided there is a morphism $i\colon S \to T$ such that $m = n \circ i$ (note that i is necessarily an embedding). This yields a preorder on the class of all embeddings with codomain X. The symmetrization $\sim$ of $\trianglelefteq$ can be characterised as follows: $m \sim n$ if, and only if, there exists an isomorphism $i\colon S \to T$ such that $m = n \circ i$. Let $\mathbb{S}\,X$ be the class of $\sim$-equivalence classes of embeddings with codomain X, equipped with the natural partial order $\leq$ induced by $\trianglelefteq$. We shall systematically represent a $\sim$-equivalence class by any of its representatives. Because $\mathcal{C}$ is well-powered and $\mathcal{M} \subseteq \{\text{monos}\}$, we see that $\mathbb{S}\,X$ is a set.

For any morphism $f\colon X \to Y$ and embedding $m\colon S \rightarrowtail X$, we can consider the $(\mathcal{Q}, \mathcal{M})$-factorisation $S \twoheadrightarrow \exists_f S \rightarrowtail Y$ of $f \circ m$. This yields a monotone map $\exists_f\colon \mathbb{S}\,X \to \mathbb{S}\,Y$ sending m to the embedding $\exists_f S \rightarrowtail Y$. (Note that the map $\exists_f$ is well-defined because factorisations are unique up to isomorphism.) Further, if $(\mathcal{Q}, \mathcal{M})$ is stable and f is a quotient, we let $f^*\colon \mathbb{S}\,Y \to \mathbb{S}\,X$ be the monotone map sending $n\colon T \rightarrowtail Y$ to its pullback along f. It is not difficult to see that f^* is right adjoint to $\exists_f$.

► **Lemma 6.** *Let $\mathcal{C}$ be any category equipped with a stable proper factorisation system, and let $f\colon X \to Y$ be any morphism in $\mathcal{C}$. The following statements hold:*
(a) *If f is an embedding, then $\exists_f\colon \mathbb{S}\,X \to \mathbb{S}\,Y$ is an order-embedding.*
(b) *If f is a quotient, then $f^*\colon \mathbb{S}\,Y \to \mathbb{S}\,X$ is an order-embedding.*

Proof. For item (a) note that, as $f\colon X \to Y$ is an embedding, $\exists_f\colon \mathbb{S}\,X \to \mathbb{S}\,Y$ sends m to $f \circ m$. Let $m_1\colon S_1 \rightarrowtail X$ and $m_2\colon S_2 \rightarrowtail X$ be embeddings such that $f \circ m_1 \leq f \circ m_2$. Then there exists $k\colon S_1 \to S_2$ such that $f \circ m_1 = f \circ m_2 \circ k$. Because f is a monomorphism, it follows that $m_1 = m_2 \circ k$, i.e. $m_1 \leq m_2$. Hence, $\exists_f$ is an order-embedding.

For item (b), it is enough to prove that $\exists_f f^* n = n$ for any $n\colon T \rightarrowtail Y$, for then $f^* n_1 \leq f^* n_2$ implies $n_1 = \exists_f f^* n_1 \leq \exists_f f^* n_2 = n_2$. Consider the pullback of f along n, as displayed on the left-hand side below.

$$
\begin{tikzcd}
f^*T \arrow[r, two heads] \arrow[d, tail, "f^*n"'] \arrow[dr, phantom, "\lrcorner", very near start] & T \arrow[d, tail, "n"] \\
X \arrow[r, two heads, "f"'] & Y
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
f^*T \arrow[r, two heads] \arrow[d, two heads] & T \arrow[d, tail, "n"] \arrow[dl, dashed] \\
\exists_f f^*T \arrow[r, tail, "\exists_f f^* n"] & Y
\end{tikzcd}
$$

Since the square on the right-hand side above commutes, there exists a diagonal filler $T \to \exists_f f^* T$. Note that this diagonal filler must be both a quotient and an embedding, hence an isomorphism. Therefore, $\exists_f f^* n = n$ in $\mathbb{S}\,Y$. ◄

3 Path Categories

3.1 Paths

Throughout this section, we fix a category $\mathcal{C}$ equipped a stable proper factorisation system.

If $(P, \leq)$ is a poset, then $C \subseteq P$ is a *chain* if it is linearly ordered. $(P, \leq)$ is a *forest* if, for all $x \in P$, the set $\downarrow x := \{y \in P \mid y \leq x\}$ is a finite chain. The *height* of a forest is the supremum of the cardinalities of its chains. The *covering relation* $\prec$ associated with a partial order $\leq$ is defined by $u \prec v$ if and only if $u < v$ and there is no w such that $u < w < v$. It is convenient to allow the empty forest. The *roots* of a forest are the minimal elements. A *tree* is a forest with at most one root. Morphisms of forests are maps which preserve roots and the covering relation. Equivalently, a monotone map $\varphi\colon U \to V$ between forests is a forest morphism if, for all $u \in U$, the restriction of φ yields a bijection between ${\downarrow}u$ and ${\downarrow}\varphi(u)$. The category of forests is denoted by $\mathcal{F}$, and the full subcategory of trees by $\mathcal{T}$. We equip these categories with the factorisation system (surjective morphisms, injective morphisms).

▶ **Definition 7.** *An object X of $\mathcal{C}$ is called a* path *provided the poset $\mathbb{S}\,X$ is a finite chain. Paths will be denoted by $P, Q, R, \ldots$.*

▶ **Example 8.** The paths in $\mathcal{F}$ and $\mathcal{T}$ are the finite chains, i.e. the trees consisting of a single branch.

▶ **Example 9.** We define a *forest-ordered σ-structure* $(\mathcal{A}, \leq)$ to be a σ-structure $\mathcal{A}$ with a forest order $\leq$ on A. A morphism of forest-ordered σ-structures $f\colon (\mathcal{A}, \leq) \to (\mathcal{B}, \leq')$ is a σ-homomorphism $f\colon \mathcal{A} \to \mathcal{B}$ that is also a forest morphism. This determines a category $\mathcal{R}(\sigma)$. We equip $\mathcal{R}(\sigma)$ with the factorisation system given by (surjective morphisms, embeddings), where an embedding is a morphism which is an embedding *qua* σ-homomorphism.

In [6], it is shown that the categories of coalgebras for the various comonads studied there are given, up to isomorphism, by subcategories of $\mathcal{R}(\sigma)$ (or minor variants thereof):

- For the Ehrenfeucht-Fraïssé comonad, this is the full subcategory $\mathcal{R}^E(\sigma)$ determined by those objects satisfying the condition (E): adjacent elements of the Gaifman graph of $\mathcal{A}$ are comparable in the forest order. For each $k > 0$, $\mathcal{R}^E_k(\sigma)$ is the full subcategory of $\mathcal{R}^E(\sigma)$ of those forest orders of height $\leq k$. The objects $(\mathcal{A}, \leq)$ of $\mathcal{R}^E_k(\sigma)$ are forest covers of $\mathcal{A}$ witnessing that its tree-depth is $\leq k$ [17].
- For the pebbling comonad, for each $k > 0$ this is the category $\mathcal{R}^P_k$ whose objects have the form $(\mathcal{A}, \leq, p)$, where $(\mathcal{A}, \leq)$ is a forest-ordered σ-structure, and $p\colon A \to [k]$ is a pebbling function. In addition to condition (E), these structures have to satisfy the condition (P): if a is adjacent to b in the Gaifman graph of $\mathcal{A}$, and $a < b$ in the forest order, then for all x such that $a < x \leq b$, $p(a) \neq p(x)$. It is shown in [6] that these structures are equivalent to the more familiar form of tree decomposition used to define tree-width [14]. Morphisms have to preserve the pebbling function.
- For the modal comonad, the category $\mathcal{R}^M_k$ has as objects the tree-ordered σ-structures of height $\leq k$ satisfying the condition (M): for $x, y \in A$, $x \prec y$ if and only if for some unique binary relation R_α in σ (“transition relation”), $R^{\mathcal{A}}_\alpha(x, y)$.

The paths in each of these categories are those structures in which the order is a finite chain. These are our key motivating examples for paths. Note that in the (multi-)modal case, ignoring propositional variables, these correspond to synchronization trees consisting of a single branch, i.e. traces.

The following fact is an immediate consequence of Lemma 6:

▶ **Lemma 10.** *Let $f\colon X \to Y$ be any morphism in $\mathcal{C}$. The following statements hold:*
(a) *If Y is a path and f is an embedding, then X is a path.*
(b) *If X is path and f is a quotient, then Y is a path.*

A *path embedding* is an embedding $P \rightarrowtail X$ whose domain is a path. Given any object X of $\mathcal{C}$, we let $\mathbb{P}\,X$ be the sub-poset of $\mathbb{S}\,X$ consisting of the path embeddings. By Lemma 10(b), for any arrow $f\colon X \to Y$, the monotone map $\exists_f\colon \mathbb{S}\,X \to \mathbb{S}\,Y$ restricts to a monotone map

$$\mathbb{P}\,f\colon \mathbb{P}\,X \to \mathbb{P}\,Y, \quad (m\colon P \rightarrowtail X) \mapsto (\exists_f m\colon \exists_f P \rightarrowtail Y).$$

By the uniqueness up to isomorphism of factorisations, this assignment is functorial.

3.2 Path categories

▶ **Definition 11.** *A* path category *is a category* $\mathcal{C}$ *satisfying the following conditions:*

(i) $\mathcal{C}$ *has a stable proper factorisation system;*

(ii) $\mathcal{C}$ *has all coproducts of small families of paths;*

(iii) *for any paths* P, Q, R*, if a composite* $P \to Q \to R$ *is a quotient, then so is* $P \to Q$*.*

▶ Remark 12. Item (iii) above is equivalent to the following condition: For any paths P, Q, R and morphisms $f \colon P \to Q$ and $g \colon Q \to R$, if any two of f, g, and $g \circ f$ are quotients, then so is the third. Thus, we shall refer to (iii) as the *2-out-of-3 condition*.

Any path category has an initial object, obtained as the coproduct of the empty family.

▶ **Example 13.** $\mathcal{F}$ and $\mathcal{T}$ are path categories. Coproducts of forests are given by disjoint union. For trees, coproducts are given by smash sum, with the bottom elements identified. Since forest morphisms preserve height, we see that $\mathcal{F}$ and $\mathcal{T}$ satisfy the 2-out-of-3 condition. Similarly, it is not difficult to see that $\mathcal{R}(\sigma)$ and its subcategories mentioned in Example 9 are all path categories ($\mathcal{R}(\sigma)$ has an initial object because we allow empty σ-structures).

▶ **Theorem 14.** *Let* $\mathcal{C}$ *be a path category. Then the assignment* $X \mapsto \mathbb{P}X$ *induces a functor* $\mathbb{P} \colon \mathcal{C} \to \mathcal{T}$ *into the category of trees.*

To prove this theorem, we start by showing that each poset $\mathbb{P}X$ is a tree.

▶ **Lemma 15.** *Let* $\mathcal{C}$ *be a path category. For any object* X *of* $\mathcal{C}$*,* $\mathbb{P}X$ *is a non-empty tree.*

Proof. Using Lemma 6(a), it is not difficult to see that the sub-poset of $\mathbb{P}X$ consisting of those elements that are below a given $P \in \mathbb{P}X$ is isomorphic to $\mathbb{P}P$. In turn, $\mathbb{P}P \cong \mathbb{S}P$ by Lemma 10(a). Since $\mathbb{S}P$ is a finite chain, we see that $\mathbb{P}X$ is a forest. Now, let $\mathbf{0} \twoheadrightarrow \widetilde{\mathbf{0}} \overset{m}{\rightarrowtail} X$ be the (quotient, embedding) factorisation of the unique morphism $\mathbf{0} \to X$ from the initial object. We claim that $m \colon \widetilde{\mathbf{0}} \rightarrowtail X$ is the unique root of $\mathbb{P}X$.

Note that $\mathbf{0}$ is a path: just observe that any embedding $S \rightarrowtail \mathbf{0}$ admits the unique morphism $\mathbf{0} \to S$ as a right inverse, and thus is a retraction. It follows that $S \cong \mathbf{0}$, i.e., $\mathbb{S}\mathbf{0}$ is the one-element poset. In particular, $\mathbf{0}$ is a path. Thus, $\widetilde{\mathbf{0}}$ is a path by Lemma 10(b). We show that $m \colon \widetilde{\mathbf{0}} \rightarrowtail X$ is the least element of $\mathbb{P}X$. If $m' \colon P \rightarrowtail X$ is any path embedding, we have a commutative square as follows.

$$\begin{tikzcd} \mathbf{0} \arrow[r, two heads] \arrow[d] & \widetilde{\mathbf{0}} \arrow[d, tail, "m"] \\ P \arrow[r, tail, "m'"] & X \end{tikzcd}$$

Hence there exists a diagonal filler $d \colon \widetilde{\mathbf{0}} \to P$, and so $m \leq m'$ in $\mathbb{P}X$. ◀

We next show that the functor $\mathbb{P}$ sends morphisms in a path category to tree morphisms, thus completing the proof of Theorem 14.

▶ **Proposition 16.** *Let* $\mathcal{C}$ *be a path category. For any arrow* f *in* $\mathcal{C}$*,* $\mathbb{P}f$ *is a tree morphism.*

Proof. It is enough to show that, for any path embedding $m \colon P \rightarrowtail X$, the induced map $\mathbb{P}f \colon {\downarrow} m \to {\downarrow} \mathbb{P}f(m)$ is a bijection.

We start by establishing surjectivity, i.e. ${\downarrow}\mathbb{P}f(m) \subseteq \mathbb{P}f({\downarrow}m)$. Let (e, j) be the (quotient, embedding) factorisation of $f \circ m$. If $n \colon Q \rightarrowtail Y$ is a path embedding such that $n \leq \mathbb{P}f(m)$ in $\mathbb{P}Y$, there exists an embedding $k \colon Q \rightarrowtail \exists_f P$ such that the left-hand diagram below commutes. Consider the pullback of k along e, as displayed in the right-hand diagram below.

$$\begin{tikzcd} P \arrow[r, tail, "m"] \arrow[dr, two heads, "e"'] & X \arrow[r, "f"] & Y & \\ & \exists_f P \arrow[u, tail, "j"] & & Q \arrow[ll, tail, "k"] \arrow[ul, tail, "n"'] \end{tikzcd} \qquad \begin{tikzcd} R \arrow[r, two heads] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & Q \arrow[d, tail, "k"] \\ P \arrow[r, two heads, "e"] & \exists_f P \end{tikzcd}$$

Then R is a path by Lemmas 5(c) and 10(a), and the composite $i \colon R \rightarrowtail P \rightarrowtail X$ is a path embedding which is below m in the poset $\mathbb{P}P$. Further, the top horizontal arrow in the pullback square is a quotient and so, by the uniqueness up to isomorphism of factorisations, the (quotient, embedding) factorisation of $f \circ i$ is $R \twoheadrightarrow Q \overset{n}{\rightarrowtail} Y$, i.e., $\mathbb{P}f(i) = n$.

For injectivity, let $m_1 \colon P_1 \rightarrowtail X$ and $m_2 \colon P_2 \rightarrowtail X$ be path embeddings in ${\downarrow}m$. Since P is a path, m_1 and m_2 are comparable in the order of $\mathbb{P}X$. Assume without loss of generality that $m_1 \leq m_2$, i.e., there exists an embedding $k \colon P_1 \rightarrowtail P_2$ such that $m_1 = m_2 \circ k$. If $\mathbb{P}f(m_1) = \mathbb{P}f(m_2)$, there exists an isomorphism $k' \colon \exists_f P_1 \to \exists_f P_2$ making the left-hand diagram below commute.

$$\begin{tikzcd} P_1 \arrow[r, two heads] \arrow[d, tail, "m_1"'] \arrow[dd, bend right=60, "k"'] & \exists_f P_1 \arrow[d, tail] \arrow[dd, bend left=60, "k'"] \\ X \arrow[r, "f"] & Y \\ P_2 \arrow[r, two heads] \arrow[u, tail, "m_2"] & \exists_f P_2 \arrow[u, tail] \end{tikzcd} \qquad \begin{tikzcd} P_1 \arrow[r, tail, "k"] \arrow[rr, bend left, two heads] & P_2 \arrow[r, two heads] & \exists_f P_2 \end{tikzcd}$$

In particular, the diagram on the right-hand side above commutes, where the top horizontal arrow is the composition of $P_1 \twoheadrightarrow \exists_f P_1$ with the isomorphism $k' \colon \exists_f P_1 \to \exists_f P_2$. By the 2-out-of-3 condition, k is an isomorphism, and so $m_1 = m_2$ in $\mathbb{P}X$. ◀

We conclude this section with the following useful observation:

▶ **Lemma 17.** *The following statements hold for any object X of a path category $\mathcal{C}$:*

(a) *Any subset $\mathcal{U} \subseteq \mathbb{P}X$ admits a supremum $\bigvee \mathcal{U}$ in $\mathbb{S}X$.*

(b) *For any path embedding $m \in \mathbb{P}X$ and non-empty set $\mathcal{S} \subseteq \mathbb{S}X$, if $m = \bigvee \mathcal{S}$ then $m \in \mathcal{S}$.*

Proof. For item (a), consider a set of path embeddings

$$\mathcal{U} = \{m_i \colon P_i \rightarrowtail X \mid i \in I\} \subseteq \mathbb{P}X.$$

Let $S \coloneqq \coprod_{i \in I} P_i$ be the coproduct in $\mathcal{C}$ of the paths P_i and consider the (quotient, embedding) factorisation of the canonical morphism $\delta \colon S \to X$ whose component at P_i is m_i:

$$\begin{tikzcd} S \arrow[r, two heads, "e"] \arrow[rr, bend left, "\delta"] & T \arrow[r, tail, "m"] & X \end{tikzcd}$$

Each path embedding $m_i \in \mathcal{U}$ factors through m, thus m is an upper bound for $\mathcal{U}$. We claim that m is the least upper bound, i.e., $m = \bigvee \mathcal{U}$ in $\mathbb{S}X$. Suppose that all path embeddings in $\mathcal{U}$ factor through some embedding $m' \colon T' \rightarrowtail X$. By the universal property of S, we get a morphism $\varphi \colon S \to T'$. Further, using again the universal property of S, it is not difficult to see that $m' \circ \varphi$ coincides with δ, and so the following square commutes.

$$\begin{tikzcd} S \arrow[r, two heads, "e"] \arrow[d, "\varphi"'] & T \arrow[d, tail, "m"] \\ T' \arrow[r, tail, "m'"] & X \end{tikzcd}$$

Therefore, there exists a diagonal filler $T \to T'$. In particular, the commutativity of the lower triangle entails that $m \leq m'$, as was to be proved.

For item (b), let $m\colon P \rightarrowtail X$ be a path embedding and $\mathcal{S} \subseteq \mathbb{S}\,X$ a non-empty set such that $m = \bigvee \mathcal{S}$. Then $n \leq m$ for each $n \in \mathcal{S}$. Since P is a path, $\downarrow m$ is a finite chain in $\mathbb{S}\,X$ and so $\mathcal{S}$ must be a finite set whose largest element coincides with m. In particular, $m \in \mathcal{S}$. ◀

4 Pathwise Embeddings, Open Maps, and Bisimulations

Throughout this section, we fix a category $\mathcal{C}$ equipped with a stable proper factorisation system.

Pathwise embeddings and open maps

Following [6], let us say that a morphism $f\colon X \to Y$ in $\mathcal{C}$ is a *pathwise embedding* if, for all path embeddings $m\colon P \rightarrowtail X$, the composite $f \circ m$ is a path embedding. Hence, $\mathbb{P}\,f(m) = f \circ m$ for all $m \in \mathbb{P}\,X$. Following again [6], we introduce a notion of open map – inspired by [13] – that, combined with the concept of pathwise embedding, will allow us to define an appropriate notion of bisimulation. A morphism $f\colon X \to Y$ in $\mathcal{C}$ is said to be *open* if it satisfies the following path-lifting property: Given any commutative square

$$\begin{tikzcd} P \arrow[r, tail] \arrow[d, tail] & Q \arrow[d, tail] \arrow[dl, dashed] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

with P, Q paths, there exists a diagonal filler $Q \to X$ (i.e., an arrow $Q \to X$ making the two triangles commute). Note that, if it exists, such a diagonal filler must be an embedding.

▶ Remark 18. The previous definition of open map differs from the one given in [13] because we require that, in the square above, the top horizontal morphism and the vertical ones be embeddings. However, a pathwise embedding is open in $\mathcal{C}$ (according to the definition above) if, and only if, it is open (in the sense of [13]) in the subcategory $\mathcal{C}_*$ of $\mathcal{C}$ having the same objects as $\mathcal{C}$ and morphisms the pathwise embeddings.

For pathwise embeddings $f\colon X \to Y$, openness can be characterised in terms of the corresponding monotone map $\mathbb{P}\,f\colon \mathbb{P}\,X \to \mathbb{P}\,Y$:

▶ **Proposition 19.** *The following are equivalent for any pathwise embedding $f\colon X \to Y$:*
1. *f is open.*
2. *$\mathbb{P}\,f$ is a p-morphism, i.e. $\mathbb{P}\,f({\uparrow} m) = {\uparrow}\,\mathbb{P}\,f(m)$ for all $m \in \mathbb{P}\,X$.*

Proof. (1) ⇒ (2). Suppose f is open, and let $m\colon P \rightarrowtail X$ be an arbitrary element of $\mathbb{P}\,X$. The inclusion $\mathbb{P}\,f({\uparrow} m) \subseteq {\uparrow}\,\mathbb{P}\,f(m)$ follows at once from monotonicity of $\mathbb{P}\,f$. For the converse inclusion, assume that $n\colon Q \rightarrowtail Y$ is an element of $\mathbb{P}\,Y$ above $\mathbb{P}\,f(m) = f \circ m$. Then, the composite $f \circ m$ must factor through n, say $f \circ m = n \circ s$ for some embedding $s\colon P \rightarrowtail Q$. Hence, we have a commutative square as displayed below.

$$\begin{tikzcd} P \arrow[r, tail, "s"] \arrow[d, tail, "m"'] & Q \arrow[d, tail, "n"] \arrow[dl, tail, dashed, "m'"'] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

Since f is open, there exists a diagonal filler $m' \colon Q \rightarrowtail X$. The commutativity of the upper triangle entails that $m' \in \uparrow m$, while the commutativity of the lower triangle implies that $\mathbb{P}\, f(m') = n$. Therefore, $\uparrow \mathbb{P}\, f(m) \subseteq \mathbb{P}\, f(\uparrow m)$.

(2) $\Rightarrow$ (1). Assume $\mathbb{P}\, f$ is a p-morphism, and consider a commutative square as follows,

$$\begin{tikzcd} P \arrow[r, tail, "s"] \arrow[d, tail, "m"'] & Q \arrow[d, tail, "n"] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

where P and Q are paths. We have $\mathbb{P}\, f(m) \leq n$ in $\mathbb{P}\, Y$, and thus there exists a path embedding $m' \colon P' \rightarrowtail X$ satisfying $m \leq m'$ and $\mathbb{P}\, f(m') = n$. The inequality $m \leq m'$ amounts to saying that $m = m' \circ l$ for some embedding $l \colon P \rightarrowtail P'$, while the equality $\mathbb{P}\, f(m') = n$ means that $f \circ m' = n \circ k$ for some isomorphism $k \colon P' \to Q$. We have a commutative diagram as displayed below.

$$\begin{tikzcd} & X \arrow[r, "f"] & Y \\ P \arrow[ur, tail, "m"] \arrow[r, tail, "l"] \arrow[rr, tail, bend right, "s"'] & P' \arrow[u, tail, "m'"'] \arrow[r, "k"] & Q \arrow[u, tail, "n"'] \end{tikzcd}$$

We claim that $m' \circ k^{-1} \colon Q \to X$ satisfies $m' \circ k^{-1} \circ s = m$ and $f \circ m' \circ k^{-1} = n$, thus showing that f is open. To start with, note that $k \circ l = s$. Just observe that

$$n \circ k \circ l = f \circ m' \circ l = f \circ m = n \circ s,$$

and so $k \circ l = s$ because n is a monomorphism. Now, by diagram chasing we see that

$$m' \circ k^{-1} \circ s = m' \circ k^{-1} \circ k \circ l = m' \circ l = m$$

and $f \circ m' \circ k^{-1} = n \circ k \circ k^{-1} = n$. This concludes the proof. ◀

Bisimulations

A *bisimulation* between objects X, Y of $\mathcal{C}$ is a span of open pathwise embeddings $X \leftarrow Z \to Y$ in $\mathcal{C}$. If such a bisimulation exists, we say that X and Y are *bisimilar*.

▶ **Example 20.** This definition directly generalizes that in [6], and the notions of bisimulation given there for the Ehrenfeucht-Fraïssé, pebbling and modal comonads are the special cases arising in the categories $\mathcal{R}_k^E(\sigma)$, $\mathcal{R}_k^P(\sigma)$ and $\mathcal{R}_k^M(\sigma)$ respectively, as described in Example 9.

▶ Remark 21. Let $\mathcal{C}$ be a path category. If we regard trees as Kripke models where the accessibility relation is the tree order, then it follows from Theorem 14 and Proposition 19 that a span of pathwise embeddings $X \xleftarrow{f} Z \xrightarrow{g} Y$ in $\mathcal{C}$ is a bisimulation if, and only if, $\mathbb{P}\, X \xleftarrow{\mathbb{P}\, f} \mathbb{P}\, Z \xrightarrow{\mathbb{P}\, g} \mathbb{P}\, Y$ is a bisimulation of Kripke models in the usual sense.

Given a bisimulation $X \leftarrow Z \rightarrow Y$, we would like to think of Z as providing a winning strategy for Duplicator in an appropriate game played "between X and Y". To substantiate this idea, in the next two sections we introduce *arboreal categories* – a refinement of the concept of path category – and show that, in these categories, bisimilarity is captured by back-and forth systems which model the dynamic nature of games.

5 Arboreal Categories

By Theorem 14, any path category $\mathcal{C}$ admits a functor $\mathbb{P}\colon \mathcal{C} \to \mathcal{T}$ into the category of trees. In general, the tree $\mathbb{P}X$ may retain little information about X. We are interested in the case where X is determined by $\mathbb{P}X$. This leads us to the notion of path-generated object.

Path-generated objects

Let $\mathcal{C}$ be a path category. For any object X of $\mathcal{C}$, we have a diagram with vertex X consisting of all path embeddings with codomain X:

$$\begin{tikzcd} & X & \\ P \arrow[ur, tail] \arrow[rr, tail] & & Q \arrow[ul, tail] \end{tikzcd}$$

The morphisms between paths are those which make the obvious triangles commute. Choosing representatives in an appropriate way, this can be seen as a cocone over the small diagram $\mathbb{P}X$. We say that X is *path-generated* provided this is a colimit cocone in $\mathcal{C}$. Intuitively, an object is path-generated if it is the colimit of its paths.

Let $\mathcal{C}_p$ be the full subcategory of $\mathcal{C}$ defined by the paths and recall that a functor $J\colon \mathcal{A} \to \mathcal{B}$ is *dense* if every $b \in \mathcal{B}$ is the colimit of the diagram $J{\downarrow}b \xrightarrow{\pi} \mathcal{A} \xrightarrow{J} \mathcal{B}$, where $J{\downarrow}b$ is the comma category and π is the natural forgetful functor.

▶ **Lemma 22.** *The following statements are equivalent for any path category $\mathcal{C}$:*
1. *Every object of $\mathcal{C}$ is path-generated.*
2. *The inclusion $\mathcal{C}_p \hookrightarrow \mathcal{C}$ is dense.*

Arboreal categories

We now state the axioms for an arboreal category. To this end, let us say that an object X of a path category $\mathcal{C}$ is *connected* if, for all small families of paths $\{P_i \mid i \in I\}$ in $\mathcal{C}$, any morphism $X \to \coprod_{i \in I} P_i$ factors through some coproduct injection $P_j \to \coprod_{i \in I} P_i$.

▶ **Definition 23.** *An* arboreal category *is a path category $\mathcal{C}$ satisfying the following conditions:*
- **(i)** *every object of $\mathcal{C}$ is path-generated;*
- **(ii)** *every path in $\mathcal{C}$ is connected.*

▶ **Example 24.** The category $\mathcal{T}$ of trees is arboreal; this is essentially the observation that (i) every tree is the colimit of the diagram given by its branches and the embeddings between them, and (ii) finite chains are connected in $\mathcal{T}$. Similarly, $\mathcal{F}$ is arboreal. Our key examples of the categories $\mathcal{R}^E_k(\sigma)$, $\mathcal{R}^P_k(\sigma)$ and $\mathcal{R}^M_k(\sigma)$ from Example 9 are also arboreal.

Note that, in view of Theorem 14, any arboreal category $\mathcal{C}$ admits a functor $\mathbb{P}\colon \mathcal{C} \to \mathcal{T}$ into the category of trees. This crucial fact is what will allow us, given an arboreal cover (cf. Section 7), to regard process structures as tree-like objects.

We collect some useful consequences of the axioms above.

▶ **Lemma 25.** *Let $\mathcal{C}$ be an arboreal category. The following statements hold:*
(a) *Between any two paths there is at most one embedding.*
(b) *For any object X of $\mathcal{C}$ and any $m \in \mathbb{S}\,X$, $m = \bigvee \{p \in \mathbb{P}\,X \mid p \leq m\}$.*
(c) *If f is a quotient in $\mathcal{C}$, then $\mathbb{P}\,f$ is a surjection.*

Proof. For item (a), since there is at most on tree morphism between any two finite chains, it suffices to show that $\mathbb{P}\colon\ \mathcal{C} \to \mathcal{T}$ is faithful on embeddings between paths. That is, whenever $f, g\colon P \rightrightarrows Q$ are embeddings between paths, $\mathbb{P}\,f = \mathbb{P}\,g$ implies $f = g$. We show that, in fact, $\mathbb{P}$ is faithful on all pathwise embeddings in $\mathcal{C}$. Suppose $f, g\colon X \rightrightarrows Y$ are pathwise embeddings. If $\mathbb{P}\,f = \mathbb{P}\,g$ then, for all path embeddings $m\colon P \rightarrowtail X$,

$$f \circ m = \mathbb{P}\,f(m) = \mathbb{P}\,g(m) = g \circ m.$$

As X is path-generated, it follows that $f = g$.

For item (b), let $m\colon S \rightarrowtail X$ be an arbitrary embedding. Clearly, $\bigvee \{p \in \mathbb{P}\,X \mid p \leq m\} \leq m$. For the converse direction, assume that $n\colon T \rightarrowtail X$ is an upper bound for $\{p \in \mathbb{P}\,X \mid p \leq m\}$. This means that each path embedding $P \rightarrowtail X$ that factors through m must factor through n. We then have a commutative diagram as displayed below.

$$\begin{tikzcd}
& T & \\
& S \arrow[u, dashed, "\gamma"'] & \\
P_i \arrow[ur, tail] \arrow[uur, bend left, tail] \arrow[rr, tail] & & P_j \arrow[ul, tail] \arrow[uul, bend right, tail]
\end{tikzcd}$$

Because S is path-generated, the cocone with vertex S is a colimit cocone. Therefore, there exists a unique mediating arrow $\gamma\colon S \to T$ making the diagram commute. Using the universal property of S, it is not difficult to see that the composite $S \xrightarrow{\gamma} T \overset{n}{\rightarrowtail} X$ coincides with m, and so $m \leq n$.

For item (c), suppose that $f\colon X \twoheadrightarrow Y$ is a quotient in $\mathcal{C}$. We first assume that $Y = P$ is a path, and then settle the general case. To show that $\mathbb{P}\,f$ is a surjection, it suffices to prove that $\mathrm{id}_P \in \mathbb{P}\,f(\mathbb{P}\,X)$, where $\mathrm{id}_P\colon P \to P$ is the identity. We have

$$\begin{aligned}
f^*\mathrm{id}_P &= \bigvee \{p \in \mathbb{P}\,X \mid p \leq f^*\mathrm{id}_P\} && \text{by item (b)} \\
&= \bigvee \{p \in \mathbb{P}\,X \mid \exists_f p \leq \mathrm{id}_P\} && \text{since } \exists_f \dashv f^* \\
&= \bigvee \mathbb{P}\,X,
\end{aligned}$$

and so (using the fact that left adjoints preserve suprema)

$$\mathrm{id}_P = \exists_f f^*\mathrm{id}_P = \bigvee \mathbb{P}\,f(\mathbb{P}\,X),$$

where the first step follows from the fact that $\exists_f \circ f^*$ is the identity of $\mathbb{S}\,P$ (cf. the proof of Lemma 6(b)). Hence, $\mathrm{id}_P \in \mathbb{P}\,f(\mathbb{P}\,X)$ by Lemma 17(b).

For the general case, let $m\colon P \rightarrowtail Y$ be an arbitrary path embedding and consider the following pullback square in $\mathcal{C}$.

$$\begin{tikzcd}
f^*P \arrow[r, two heads, "g"] \arrow[d, tail, "f^*m"'] \arrow[dr, phantom, "\lrcorner", very near start] & P \arrow[d, tail, "m"] \\
X \arrow[r, two heads, "f"] & Y
\end{tikzcd}$$

By the argument above, there is a path embedding $n\colon Q \rightarrowtail f^*P$ such that $\mathbb{P}\,g(n) = \mathrm{id}_P$. It follows easily that $\mathbb{P}\,f(f^*m \circ n) = m$. ◀

The next proposition will allow us to construct an object from a prescribed set of path embeddings, without adding any new paths in the process.

▶ **Proposition 26.** *Let $\mathscr{C}$ be an arboreal category, X an object of $\mathscr{C}$, and $\mathcal{U} \subseteq \mathbb{P}X$. A path embedding $m \in \mathbb{P}X$ is below $\bigvee \mathcal{U}$ if, and only if, it is below some element of $\mathcal{U}$.*

Proof. Fix an arbitrary object X of $\mathscr{C}$ and a set of path embeddings $\mathcal{U} = \{m_i \colon P_i \rightarrowtail X \mid i \in I\}$. Let $m \colon P \rightarrowtail X$ be an arbitrary path embedding. If m is below some element of $\mathcal{U}$, then clearly $m \leq \bigvee \mathcal{U}$. For the converse direction, suppose that $m \leq \bigvee \mathcal{U}$. Recall from the proof of Lemma 17(a) that the supremum of $\mathcal{U}$ is obtained by taking the (quotient, embedding) factorisation $\coprod_{i \in I} P_i \overset{e}{\twoheadrightarrow} S \overset{n}{\rightarrowtail} X$ of the canonical morphism $\coprod_{i \in I} P_i \to X$. With this notation, $\bigvee \mathcal{U} = n$. Since $m \leq \bigvee \mathcal{U}$, there exists an embedding $m' \colon P \rightarrowtail S$ such that $m = n \circ m'$. Consider the pullback of m' along e:

$$
\begin{tikzcd}
T \arrow[r, two heads, "r"] \arrow[d, tail, "j"'] \arrow[dr, phantom, "\lrcorner", very near start] & P \arrow[d, tail, "m'"] \\
\coprod_{i \in I} P_i \arrow[r, two heads, "e"] & S
\end{tikzcd}
$$

Applying Lemma 25(c) to the quotient r, we see that there exists a path embedding $k \colon Q \rightarrowtail T$ such that $r \circ k$ is a quotient. Because Q is connected, $j \circ k \colon Q \rightarrowtail \coprod_{i \in I} P_i$ must factor through some coproduct injection $\varphi_i \colon P_i \to \coprod_{i \in I} P_i$, i.e., $j \circ k = \varphi_i \circ p$ for some path embedding $p \colon Q \rightarrowtail P$. We then have a commutative diagram as follows.

$$
\begin{tikzcd}
Q \arrow[r, two heads, "r \circ k"] \arrow[d, tail, "p"'] & P \arrow[d, tail, "m'"'] \arrow[dr, tail, "m"] & \\
P_i \arrow[r, "e \circ \varphi_i"] \arrow[rr, tail, bend right, "m_i"'] & S \arrow[r, tail, "n"] & X
\end{tikzcd}
$$

As $m \circ r \circ k = m_i \circ p$ and the right-hand side of the equation is an embedding, $r \circ k$ is an isomorphism. So $m \leq m_i \in \mathcal{U}$, thus concluding the proof. ◀

6 Back-and-Forth Systems and Games

Throughout this section, we work in a fixed arboreal category $\mathscr{C}$. First, we introduce back-and-forth systems in $\mathscr{C}$ and show that they capture precisely the bisimilarity relation defined in Section 4 in terms of spans of open pathwise embeddings. Then, we show that back-and-forth systems can be equivalently seen as appropriate back-and-forth games.

Back-and-forth systems

Given objects X and Y of $\mathscr{C}$, we consider spans of (equivalence classes of) path embeddings of the form $X \overset{m}{\leftarrowtail} P \overset{n}{\rightarrowtail} Y$. Such a span can be thought of as a partial isomorphism "of shape P" between X and Y. A back-and-forth system between X and Y is a collection of such spans containing an "initial element" and satisfying an appropriate extension property.

Let X, Y be any two objects of $\mathscr{C}$. Given $m \in \mathbb{P}X$ and $n \in \mathbb{P}Y$, we write $[\![m, n]\!]$ to denote that $\mathrm{dom}(m) \cong \mathrm{dom}(n)$. Observe that (i) any two embeddings in the same $\sim$-equivalence class have isomorphic domains, and (ii) given $[\![m, n]\!]$, there exist path embeddings $m' \sim m$ and $n' \sim n$ such that $\mathrm{dom}(m') = \mathrm{dom}(n')$. Hence, the pairs of the form $[\![m, n]\!]$ capture the partial isomorphisms $X \overset{m}{\leftarrowtail} P \overset{n}{\rightarrowtail} Y$ "of shape P".

▶ **Definition 27.** *A* back-and-forth system *between objects X, Y of $\mathcal{C}$ is a set $\mathcal{B} = \{[\![m_i, n_i]\!] \mid m_i \in \mathbb{P}\, X,\ n_i \in \mathbb{P}\, Y,\ i \in I\}$ satisfying the following conditions:*

(i) $[\![\bot_X, \bot_Y]\!] \in \mathcal{B}$*, where $\bot_X$ and $\bot_Y$ are the least elements of $\mathbb{P}\, X$ and $\mathbb{P}\, Y$, respectively;*

(ii) *if $[\![m, n]\!] \in \mathcal{B}$ and $m' \in \mathbb{P}\, X$ are such that $m \prec m'$, there exists $n' \in \mathbb{P}\, Y$ satisfying $n \prec n'$ and $[\![m', n']\!] \in \mathcal{B}$;*

(iii) *if $[\![m, n]\!] \in \mathcal{B}$ and $n' \in \mathbb{P}\, Y$ are such that $n \prec n'$, there exists $m' \in \mathbb{P}\, X$ satisfying $m \prec m'$ and $[\![m', n']\!] \in \mathcal{B}$.*

A back-and-forth system $\mathcal{B}$ is strong *if, for all $[\![m, n]\!] \in \mathcal{B}$ and all $m' \in \mathbb{P}\, X$, $n' \in \mathbb{P}\, Y$, if $m' \prec m$ and $n' \prec n$ then $[\![m', n']\!] \in \mathcal{B}$.*

Two objects X and Y of $\mathcal{C}$ are said to be *(strong) back-and-forth equivalent* if there exists a (strong) back-and-forth system between them.

▶ Remark 28. The definition of (strong) back-and-forth system given above is a variant of the notion of *(strong) path bisimulation* from [12]. The nomenclature adopted here is motivated by the analogy with back-and-forth systems of partial isomorphisms from model theory [15].

The aim of this section is to prove the following result:

▶ **Theorem 29.** *In an arboreal category with binary products, any two objects are bisimilar if, and only if, they are strong back-and-forth equivalent.*

In all our key examples of arboreal categories, binary products exist:

▶ **Example 30.** The category $\mathcal{T}$ has binary products. These are computed as "synchronous products" consisting of the pairs (x, y) of elements having the same height, with the componentwise order. Similarly, $\mathcal{F}$ has binary products, and so does the category $\mathcal{R}(\sigma)$ if the relations in the synchronous product are interpreted componentwise. As synchronous products do not increase the height of forests, we see that the category $\mathcal{R}^E_k(\sigma)$ from Example 9 has binary products. Finally, binary products exist in $\mathcal{R}^P_k(\sigma)$ and $\mathcal{R}^M_k(\sigma)$ and can be described again as variants of synchronous products.

▶ Remark 31. Direct inspection of the relevant proofs shows that Theorem 29 can be slightly generalized to the effect that, for any two objects X and Y of an arboreal category, if the product $X \times Y$ exists, then X and Y are bisimilar precisely when they are strong back-and-forth equivalent.

We start by establishing the easy direction of Theorem 29.

▶ **Proposition 32.** *Any two bisimilar objects of an arboreal category are strong back-and-forth equivalent.*

Proof. Suppose that $X \xleftarrow{f} Z \xrightarrow{g} Y$ is a span of open pathwise embeddings in an arboreal category $\mathcal{C}$. We claim that

$$\mathcal{B} \coloneqq \{[\![\mathbb{P}\, f(m), \mathbb{P}\, g(m)]\!] \mid m \in \mathbb{P}\, Z\}$$

is a strong back-and-forth system between X and Y. We show that items (i), (ii), and (iii) in Definition 27 are satisfied.

For item (i), let $\bot_Z$ be the least element of $\mathbb{P}\, Z$. Then $\mathbb{P}\, f(\bot_Z) = \bot_X$ and $\mathbb{P}\, g(\bot_Z) = \bot_Y$ because tree morphisms preserve roots, and so $[\![\bot_X, \bot_Y]\!] \in \mathcal{B}$. For item (ii), let $[\![\mathbb{P}\, f(m), \mathbb{P}\, g(m)]\!] \in \mathcal{B}$ and $m' \in \mathbb{P}\, X$ be such that $\mathbb{P}\, f(m) \prec m'$. Let us denote $P \coloneqq \mathrm{dom}(m)$ and $P' \coloneqq \mathrm{dom}(m')$. As $f \circ m \leq m'$ in $\mathbb{P}\, X$, there exists $k \colon P \rightarrowtail P'$ such that $f \circ m = m' \circ k$ in $\mathcal{C}$. Therefore, we have a commutative square as follows.

$$\begin{tikzcd} P \arrow[r, tail, "k"] \arrow[d, tail, "m"'] & P' \arrow[d, tail, "m'"] \\ Z \arrow[r, "f"] & X \end{tikzcd}$$

Since f is open, there exists a diagonal filler $n\colon P' \rightarrowtail Z$. Thus, $m' = \mathbb{P}\, f(n)$ and $[\![m', \mathbb{P}\, g(n)]\!] \in \mathcal{B}$. It remains to show that $\mathbb{P}\, g(m) \prec \mathbb{P}\, g(n)$. Note that $m \leq n$ and $\mathbb{P}\, f(m) \prec \mathbb{P}\, f(n)$ entail $m \prec n$, and so $\mathbb{P}\, g(m) \prec \mathbb{P}\, g(n)$ because $\mathbb{P}\, g$ preserves the covering relation.

The proof of item (iii) is the same, mutatis mutandis, as for (ii). Finally, observe that the back-and-forth system $\mathcal{B}$ is strong. To see this, suppose that $[\![\mathbb{P}\, f(m), \mathbb{P}\, g(m)]\!] \in \mathcal{B}$, and let $p_1 \in \mathbb{P}\, X$ and $p_2 \in \mathbb{P}\, Y$ satisfy $p_1 \prec \mathbb{P}\, f(m)$ and $p_2 \prec \mathbb{P}\, g(m)$. As $\mathbb{P}\, f$ is a tree morphism, there exists $m' \in \mathbb{P}\, Z$ such that $m' \prec m$ and $\mathbb{P}\, f(m') = p_1$. But then $\mathbb{P}\, g(m') \prec \mathbb{P}\, g(m)$ implies $\mathbb{P}\, g(m') = p_2$, and therefore $[\![p_1, p_2]\!] \in \mathcal{B}$. ◀

To establish the other direction of Theorem 29, we start by considering a strong back-and-forth system $\mathcal{B} = \{[\![m_i, n_i]\!] \mid i \in I\}$ between X and Y, and attempt to construct an object Z and a span of open pathwise embeddings $X \leftarrow Z \rightarrow Y$. Intuitively, Z is obtained by gluing together the paths $P_i \coloneqq \mathrm{dom}(m_i)$, for $i \in I$, by taking a colimit in $\mathcal{C}$. This colimit can be equivalently described as the supremum of a set of path embeddings as we now explain.

Consider an arbitrary $[\![m_i, n_i]\!] \in \mathcal{B}$ and assume without loss of generality that $\mathrm{dom}(m_i) = P_i = \mathrm{dom}(n_i)$ for some path P_i. Then the product arrow $\langle m_i, n_i \rangle\colon P_i \to X \times Y$ is an embedding. In fact, it suffices that m_i be an embedding (or, symmetrically, that n_i be an embedding), for then $m_i = \pi_X \circ \langle m_i, n_i \rangle$ entails that $\langle m_i, n_i \rangle$ is an embedding, where $\pi_X\colon X \times Y \to X$ is the projection. Therefore, we can identify each $[\![m_i, n_i]\!] \in \mathcal{B}$ with a path embedding $\langle m_i, n_i \rangle \in \mathbb{P}\,(X \times Y)$ and compute the supremum $m\colon Z \rightarrowtail X \times Y$ in $\mathbb{S}\,(X \times Y)$ of all these path embeddings. (It is not difficult to see that the assignment $[\![m_i, n_i]\!] \mapsto \langle m_i, n_i \rangle \in \mathbb{P}\,(X \times Y)$ does not depend on the choice of the representatives in the equivalence classes of m_i and n_i.) We note in passing the following immediate fact:

▶ **Lemma 33.** *Let $\mathcal{B} = \{[\![m_i, n_i]\!] \mid i \in I\}$ be a back-and-forth system between X and Y. If $\mathcal{B}$ is strong, then $\{\langle m_i, n_i \rangle \in \mathbb{P}\,(X \times Y) \mid i \in I\}$ is downwards closed in $\mathbb{S}\,(X \times Y)$.*

To show that the span $X \xleftarrow{\pi_X \circ m} Z \xrightarrow{\pi_Y \circ m} Y$ is a bisimulation, we exploit the fact that Z does not admit more path embeddings than those prescribed (cf. Proposition 26). The following proposition then completes the proof of Theorem 29:

▶ **Proposition 34.** *In an arboreal category with binary products, any two strong back-and-forth equivalent objects are bisimilar.*

Proof. Let $\mathcal{C}$ be an arboreal category with binary products, and let X, Y be any two objects of $\mathcal{C}$. Assume that there is a strong back-and-forth system $\mathcal{B} = \{[\![m_i, n_i]\!] \mid i \in I\}$ between X and Y, and consider the set

$$\mathcal{U} \coloneqq \{\langle m_i, n_i \rangle \in \mathbb{P}\,(X \times Y) \mid i \in I\}.$$

Let $m\colon Z \rightarrowtail X \times Y$ be the supremum of $\mathcal{U}$ in $\mathbb{S}\,(X \times Y)$. We claim that

$$X \xleftarrow{\pi_X \circ m} Z \xrightarrow{\pi_Y \circ m} Y$$

is a bisimulation between X and Y.

To see that this is a span of pathwise embeddings, consider an arbitrary path embedding $n\colon P \rightarrowtail Z$. In view of Proposition 26 and Lemma 33, $m \circ n \in \mathcal{U}$. That is, $m \circ n = \langle m_i, n_i \rangle$ in $\mathbb{P}(X \times Y)$ for some $[\![m_i, n_i]\!] \in \mathcal{B}$. It follows that $m \circ n = \langle m_i, n_i \rangle \circ \varphi$ in $\mathcal{C}$ for some isomorphism φ, and so $\pi_X \circ m \circ n$ and $\pi_Y \circ m \circ n$ are embeddings because $\pi_X \circ m \circ n = m_i \circ \varphi$ and $\pi_Y \circ m \circ n = n_i \circ \varphi$.

It remains to show that $\pi_X \circ m$ and $\pi_Y \circ m$ are open. We prove that $\pi_X \circ m$ is open; the proof for $\pi_Y \circ m$ follows by symmetry. Consider a commutative square in $\mathcal{C}$ as displayed below, where P and Q are paths.

$$\begin{tikzcd} P \arrow[r, tail, "k"] \arrow[d, tail, "n"'] & Q \arrow[d, tail, "m_j"] \\ Z \arrow[r, "\pi_X \circ m"] & X \end{tikzcd}$$

Reasoning as above, we see that $m \circ n = \langle m_i, n_i \rangle$ in $\mathbb{P}(X \times Y)$ for some $[\![m_i, n_i]\!] \in \mathcal{B}$. Therefore, in $\mathbb{P}X$, we have $m_i = \pi_X \circ m \circ n \leq m_j$. Applying item (ii) in Definition 27 (possibly finitely many times), it follows that there exists $n_j \in \mathbb{P}Y$ such that $n_i \leq n_j$ and $[\![m_j, n_j]\!] \in \mathcal{B}$. Suppose without loss of generality that $\mathrm{dom}(m_j) = \mathrm{dom}(n_j)$. Then, $\langle m_j, n_j \rangle \in \mathcal{U}$ and so $\langle m_j, n_j \rangle\colon Q \rightarrowtail X \times Y$ factors through the supremum $m\colon Z \rightarrowtail X \times Y$ of $\mathcal{U}$. That is, $\langle m_j, n_j \rangle = m \circ h$ in $\mathcal{C}$ for some morphism $h\colon Q \to Z$. We claim that the following diagram commutes, thus establishing that $\pi_X \circ m$ is open.

$$\begin{tikzcd} P \arrow[r, tail, "k"] \arrow[d, tail, "n"'] & Q \arrow[d, tail, "m_j"] \arrow[dl, "h"'] \\ Z \arrow[r, "\pi_X \circ m"] & X \end{tikzcd}$$

For the commutativity of the lower triangle, just observe that

$$\pi_X \circ m \circ h = \pi_X \circ \langle m_j, n_j \rangle = m_j.$$

Now, assume without loss of generality that $\mathrm{dom}(m_i) = R = \mathrm{dom}(n_i)$ for some path R. As already observed above, $m \circ n = \langle m_i, n_i \rangle$ in $\mathbb{P}(X \times Y)$ implies that $m \circ n = \langle m_i, n_i \rangle \circ \varphi$ in $\mathcal{C}$ for some isomorphism $\varphi\colon P \to R$. Thus, for the upper triangle, we have

$$\begin{aligned} n = h \circ k \;&\Leftrightarrow\; m \circ n = m \circ h \circ k \\ &\Leftrightarrow\; \langle m_i, n_i \rangle \circ \varphi = \langle m_j, n_j \rangle \circ k \\ &\Leftrightarrow\; \begin{cases} m_i \circ \varphi = m_j \circ k \\ n_i \circ \varphi = n_j \circ k \end{cases} \end{aligned}$$

where in the first step we used the fact that m is a monomorphism. In turn, the inequalities $m_i \leq m_j$ and $n_i \leq n_j$ entail the existence of embeddings $k_1, k_2\colon R \rightrightarrows Q$ such that $m_i = m_j \circ k_1$ and $n_i = n_j \circ k_2$. By Lemma 25(a) we have $k_1 \circ \varphi = k = k_2 \circ \varphi$. It follows that $m_i \circ \varphi = m_j \circ k$ and $n_i \circ \varphi = n_j \circ k$, and so $n = h \circ k$. ◀

Back-and-forth games

Let $\mathcal{C}$ be an arboreal category and let X, Y be any two objects of $\mathcal{C}$. We define a back-and-forth game $\mathcal{G}(X, Y)$ played by Spoiler and Duplicator on X and Y as follows. Positions in the game are pairs of (equivalence classes of) path embeddings $(m, n) \in \mathbb{P}X \times \mathbb{P}Y$. The winning relation $\mathcal{W}(X, Y) \subseteq \mathbb{P}X \times \mathbb{P}Y$ consists of the pairs (m, n) such that $\mathrm{dom}(m) \cong \mathrm{dom}(n)$.

Let $\bot_X : P \rightarrowtail X$ and $\bot_Y : Q \rightarrowtail Y$ be the roots of $\mathbb{P}X$ and $\mathbb{P}Y$, respectively. If $P \ncong Q$, then Duplicator loses the game. Otherwise, the initial position is $(\bot_X, \bot_Y)$. At the start of each round, the position is specified by a pair $(m, n) \in \mathbb{P}X \times \mathbb{P}Y$, and the round proceeds as follows: Either Spoiler chooses some $m' \succ m$ and Duplicator must respond with some $n' \succ n$, or Spoiler chooses some $n'' \succ n$ and Duplicator must respond with $m'' \succ m$. Duplicator wins the round if they are able to respond and the new position is in $\mathcal{W}(X, Y)$. Duplicator wins the game if they have a strategy which is winning after t rounds, for all $t \geq 0$.

▶ Remark 35. It is shown in [6] that the abstract game $\mathcal{G}(X, Y)$ restricts, in the case of the arboreal categories $\mathcal{R}^E_k(\sigma)$, $\mathcal{R}^P_k(\sigma)$ and $\mathcal{R}^M_k(\sigma)$, to the usual k-round Ehrenfeucht-Fraïssé, k-pebble and k-round bisimulation games, respectively.

The following straightforward observation makes precise the translation between strong back-and-forth systems and back-and-forth games.

▶ **Lemma 36.** *Two objects X, Y of an arboreal category $\mathcal{C}$ are strong back-and-forth equivalent if, and only if, Duplicator has a winning strategy in the game $\mathcal{G}(X, Y)$.*

Proof. Clearly, if $\mathcal{B} = \{[\![m_i, n_i]\!] \mid i \in I\}$ is a strong back-and-forth system between X and Y, then the plays in the set

$$\{(m_i, n_i) \mid i \in I\} \subseteq \mathbb{P}X \times \mathbb{P}Y$$

yield a winning strategy for Duplicator in the game $\mathcal{G}(X, Y)$.

Conversely, a winning strategy for Duplicator in the game $\mathcal{G}(X, Y)$ determines a set $W \subseteq \mathcal{W}(X, Y)$ of the plays following this strategy, and

$$\mathcal{B} \coloneqq \{[\![m, n]\!] \mid (m, n) \in W\}$$

is a back-and-forth system. It is not difficult to see that $\mathcal{B}$ is strong. ◀

The previous lemma, combined with Theorem 29, yields at once the following result:

▶ **Theorem 37.** *Let $\mathcal{C}$ be an arboreal category with binary products. Any two objects X, Y of $\mathcal{C}$ are bisimilar if, and only if, Duplicator has a winning strategy in the game $\mathcal{G}(X, Y)$.*

7 Arboreal Covers

We return to the underlying motivation for the axiomatic development in this paper. Arboreal categories have a rich intrinsic process structure, which allows "dynamic" notions such as bisimulation and back-and-forth games, and resource notions such as the height of a tree, to be defined. A key idea is to relate these process notions to extensional, or "static" structures. In particular, much of finite model theory and descriptive complexity can be seen in this way.

In the general setting, we have an arboreal category $\mathcal{C}$, and another category $\mathcal{E}$, which we think of as the extensional category.

▶ **Definition 38.** *An* arboreal cover *of $\mathcal{E}$ by $\mathcal{C}$ is given by a comonadic adjunction*

$$\begin{tikzcd} \mathcal{C} \arrow[r, bend left, "L"] \arrow[r, phantom, "\perp"] & \mathcal{E} \arrow[l, bend left, "R"] \end{tikzcd}.$$

As for any adjunction, this induces a comonad on $\mathcal{E}$. The comonad is (G, ε, δ), where $G \coloneqq LR$, ε is the counit of the adjunction, and $\delta_a : LRa \to LRLRa$ is given by $\delta_a \coloneqq L(\eta_{Ra})$, with η the unit of the adjunction. The comonadicity condition states that the Eilenberg-Moore category of coalgebras for this comonad is isomorphic to $\mathcal{C}$. The idea is then that we can use the arboreal category $\mathcal{C}$, with its rich process structure and all the associated notions, to study the extensional category $\mathcal{E}$ via the adjunction. Both the co-Kleisli category of the comonad, and the full Eilenberg-Moore category, are useful in this regard.

We now bring resources into the picture.

▶ **Definition 39.** *Let $\mathcal{C}$ be an arboreal category, with full subcategory of paths $\mathcal{C}_p$. We say that $\mathcal{C}$ is* resource-indexed *by a resource parameter k if for all $k \geq 0$, there is a full subcategory $\mathcal{C}_p^k$ of $\mathcal{C}_p$ closed under embeddings*[2] *with*

$$\mathcal{C}_p^0 \hookrightarrow \mathcal{C}_p^1 \hookrightarrow \mathcal{C}_p^2 \hookrightarrow \cdots$$

This induces a corresponding tower of full subcategories $\mathcal{C}_k$ of $\mathcal{C}$, with the objects of $\mathcal{C}_k$ those whose cocone of path embeddings with domain in $\mathcal{C}_p^k$ is a colimit cocone in $\mathcal{C}$.

▶ **Example 40.** One resource parameter which is always available is to take $\mathcal{C}_p^k$ to be given by those paths in $\mathcal{C}$ whose chain of subobjects is of length $\leq k$. In the case of $\mathcal{F}$ and $\mathcal{T}$, the corresponding categories $\mathcal{F}_k$ and $\mathcal{T}_k$ are the forests and trees of height $\leq k$. We can think of this as a temporal parameter, restricting the number of sequential steps, or the number of rounds in a game. For the Ehrenfeucht-Fraïssé and modal comonads, we recover $\mathcal{R}_k^E$ and $\mathcal{R}_k^M$ as described in Example 9, corresponding to k-round versions of the Ehrenfeucht-Fraïssé and modal bisimulation games respectively [6]. However, note that for the pebbling comonad, the relevant resource index is the number of pebbles, which is a memory restriction along a computation or play of a game. This leads to $\mathcal{R}_k^P$ as described in Example 9.

In Proposition 42 below we shall see that, given a resource-indexed arboreal category $\{\mathcal{C}_k\}$, each category $\mathcal{C}_k$ is arboreal. This allows us to exploit the ideas developed in this paper for any choice of the resource parameter k. We start by proving the following fact:

▶ **Lemma 41.** *Let $\{\mathcal{C}_k\}$ be a resource-indexed arboreal category and suppose that $X \rightarrowtail Y$ is an embedding in $\mathcal{C}$. For any k, if $Y \in \mathcal{C}_k$ then also $X \in \mathcal{C}_k$.*

Proof. We start by showing that, for any path embedding $P \rightarrowtail Y$ in $\mathcal{C}$, if $Y \in \mathcal{C}_k$ then $P \in \mathcal{C}_p^k$. Consider the set

$$\mathcal{U} \coloneqq \{p \in \mathbb{P}Y \mid \mathrm{dom}(p) \in \mathcal{C}_p^k\}.$$

(Note that $\mathcal{U}$ is well-defined because any two representatives in the equivalence class of p have isomorphic domains, and $\mathcal{C}_p^k$ is closed under isomorphisms.) As Y is the colimit in $\mathcal{C}$ of the subdiagram of $\mathbb{P}Y$ consisting of those path embeddings whose domain is in $\mathcal{C}_p^k$, it follows that $\bigvee \mathcal{U} = \mathrm{id}_Y$ in $\mathbb{S}Y$. If there exists a path embedding $m \colon P \rightarrowtail Y$, then $m \leq \bigvee \mathcal{U}$ in $\mathbb{P}Y$ and so, by Proposition 26, m factors through some $p \in \mathcal{U}$. In particular, there exists an embedding $P \rightarrowtail \mathrm{dom}(p)$. Because $\mathrm{dom}(p) \in \mathcal{C}_p^k$, we see that $P \in \mathcal{C}_p^k$.

Now, suppose that $j \colon X \rightarrowtail Y$ is an embedding in $\mathcal{C}$ and $Y \in \mathcal{C}_k$. Since X is path-generated, it is the colimit in $\mathcal{C}$ of the small diagram $\mathbb{P}X$. We show that, for any path embedding $P \rightarrowtail X$, the path P must belong to $\mathcal{C}_p^k$. It then follows immediately that $X \in \mathcal{C}_k$. Let $m \colon P \rightarrowtail X$ be an arbitrary path embedding. The composite $j \circ m \colon P \rightarrowtail Y$ is also a path embedding, and so $P \in \mathcal{C}_p^k$ by the argument above. ◀

[2] That is, for any embedding $P \rightarrowtail Q$ in $\mathcal{C}$ with P, Q paths, if $Q \in \mathcal{C}_p^k$ then also $P \in \mathcal{C}_p^k$.

▶ **Proposition 42.** *Let $\{\mathcal{C}_k\}$ be a resource-indexed arboreal category. Then $\mathcal{C}_k$ is an arboreal category for each k.*

Proof. If $\mathcal{C}$ is equipped with the stable proper factorisation system $(\mathcal{Q}, \mathcal{M})$, consider the classes of morphisms $\mathcal{Q}' \coloneqq \mathcal{Q} \cap \mathcal{C}_k$ and $\mathcal{M}' \coloneqq \mathcal{M} \cap \mathcal{C}_k$. It is not difficult to see that $(\mathcal{Q}', \mathcal{M}')$ is a proper factorisation system in $\mathcal{C}_k$. Just observe that, whenever $W \to Z$ is a morphism in $\mathcal{C}_k$ and

$$W \twoheadrightarrow X \rightarrowtail Y$$

is its (quotient, embedding) factorisation in $\mathcal{C}$, then $X \in \mathcal{C}_k$ by Lemma 41. Using again Lemma 41, along with the fact that embeddings are stable under pullbacks, it follows at once that $(\mathcal{Q}', \mathcal{M}')$ is stable (and, in fact, pullbacks of $\mathcal{Q}'$-morphisms along $\mathcal{M}'$-morphisms are computed in $\mathcal{C}$). With respect to this factorisation system, it is not difficult to see that the paths in $\mathcal{C}_k$ are precisely the objects of $\mathcal{C}_p^k$.

Moreover, $\mathcal{C}_k$ has all coproducts of small families of paths, and they are computed in $\mathcal{C}$. To see this, consider a set of paths $\{P_i \in \mathcal{C}_p^k \mid i \in I\}$ and let $\coprod_{i\in I} P_i$ be the coproduct in $\mathcal{C}$. If $m \colon P \rightarrowtail \coprod_{i\in I} P_i$ is any path embedding in $\mathcal{C}$ then, because P is connected, m must factor through some coproduct injection. In particular, there exist $i \in I$ and an embedding $P \rightarrowtail P_i$. Since $P_i \in \mathcal{C}_p^k$, we get $P \in \mathcal{C}_p^k$. As $\coprod_{i\in I} P_i$ is path-generated in $\mathcal{C}$, it follows at once that $\coprod_{i\in I} P_i \in \mathcal{C}_k$. Hence, $\coprod_{i\in I} P_i$ coincides with the coproduct of the family $\{P_i \in \mathcal{C}_p^k \mid i \in I\}$ in $\mathcal{C}_k$.

We conclude that $\mathcal{C}_k$ is a path category. Further, every object of $\mathcal{C}_k$ is path-generated by definition, and paths in $\mathcal{C}_k$ are connected because any path in $\mathcal{C}_k$ is also a path in $\mathcal{C}$ and, as observed above, coproducts of paths in $\mathcal{C}_k$ are computed in $\mathcal{C}$. Therefore, $\mathcal{C}_k$ is arboreal. ◀

▶ **Definition 43.** *Let $\{\mathcal{C}_k\}$ be a resource-indexed arboreal category. A* resource-indexed arboreal cover *of $\mathcal{E}$ by $\mathcal{C}$ is an indexed family of comonadic adjunctions*

$$\begin{tikzcd} \mathcal{C}_k \arrow[r, bend left, "L_k"] \arrow[r, phantom, "\perp"] & \mathcal{E} \arrow[l, bend left, "R_k"] \end{tikzcd}$$

with corresponding comonads G_k on $\mathcal{E}$.

▶ **Example 44.** Our key examples arise by taking the extensional category $\mathcal{E}$ to be $\mathsf{Struct}(\sigma)$. For each $k \geq 0$, there are evident forgetful functors

$$L_k^E \colon \mathcal{R}_k^E \to \mathsf{Struct}(\sigma), \quad L_k^P \colon \mathcal{R}_k^P \to \mathsf{Struct}(\sigma), \quad L_k^M \colon \mathcal{R}_k^M \to \mathsf{Struct}(\sigma)$$

which forget the forest order, and in the case of $\mathcal{R}_k^P$, also the pebbling function. These functors are all comonadic over $\mathsf{Struct}(\sigma)$. (To be precise, in the modal logic case the category $\mathcal{E}$ consists of *pointed* structures, cf. [6].) The right adjoints build a forest over a structure $\mathcal{A}$ by forming sequences of elements over the universe A, suitably labelled and with the σ-relations interpreted so as to satisfy the conditions (E), (P) and (M) respectively. This yields the comonads described concretely in [1, 6]. The sequences correspond to plays in the Ehrenfeucht-Fraïssé, pebbling and modal bisimulation games respectively.

We now show how resource-indexed arboreal covers can be used to define important notions on the extensional category. For a resource-indexed arboreal cover of $\mathcal{E}$ by $\mathcal{C}$, with adjunctions L_k, R_k and comonads G_k, we define three resource-indexed equivalence relations on objects of $\mathcal{E}$. The first two use the co-Kleisli category of G_k, while the third uses the Eilenberg-Moore category.

▶ **Definition 45.** *Consider a resource-indexed arboreal cover of $\mathcal{E}$ by $\mathcal{C}$, and objects $a, b \in \mathcal{E}$.*
1. *We define $a \rightleftarrows_k^{\mathcal{C}} b$ iff there are co-Kleisli morphisms $G_k a \to b$ and $G_k b \to a$.*
2. *We define $a \cong_k^{\mathcal{C}} b$ iff a and b are isomorphic in the co-Kleisli category of G_k.*
3. *We define $a \leftrightarrow_k^{\mathcal{C}} b$ iff there is a bisimulation between $R_k a$ and $R_k b$ in $\mathcal{C}_k$.*

▶ **Proposition 46.** *Assume $\mathcal{E}$ has binary products. For objects a and b of $\mathcal{E}$, $a \leftrightarrow_k^{\mathcal{C}} b$ iff $R_k a$ and $R_k b$ are strong back-and-forth equivalent, iff Duplicator has a winning strategy in the game $\mathcal{G}(R_k a, R_k b)$.*

Proof. This follows directly from Theorems 29 and 37, and Proposition 42 (cf. also Remark 31). Just observe that, since right adjoints preserve limits, the product of $R_k a$ and $R_k b$ in $\mathcal{C}_k$ exists and can be identified with $R_k(a \times b)$. ◀

What do these notions mean in Example 44? For each of our three types of model comparison game, there are corresponding fragments $\mathcal{L}_k$ of first-order logic [15, 8]:
- For Ehrenfeucht-Fraïssé games, $\mathcal{L}_k$ is the fragment of quantifier-rank $\leq k$.
- For pebble games, $\mathcal{L}_k$ is the k-variable fragment.
- For bisimulation games, $\mathcal{L}_k$ is the modal fragment with modal depth $\leq k$.

In each case, we write $\exists\mathcal{L}_k$ for the existential positive fragment of $\mathcal{L}_k$, and $\mathcal{L}_k(\#)$ for the extension of $\mathcal{L}_k$ with counting quantifiers [15]. For each logic $\mathcal{L}$, there is the usual equivalence on σ-structures: $\mathcal{A} \equiv^{\mathcal{L}} \mathcal{B}$ iff for all φ in $\mathcal{L}$, $\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi$.

We now have the following result from [6]:

▶ **Theorem 47.** *For σ-structures $\mathcal{A}$ and $\mathcal{B}$:*
1. $\mathcal{A} \equiv^{\exists\mathcal{L}_k} \mathcal{B} \iff \mathcal{A} \rightleftarrows_k^{\mathcal{C}} \mathcal{B}$.
2. $\mathcal{A} \equiv^{\mathcal{L}_k} \mathcal{B} \iff \mathcal{A} \leftrightarrow_k^{\mathcal{C}} \mathcal{B}$.
3. $\mathcal{A} \equiv^{\mathcal{L}_k(\#)} \mathcal{B} \iff \mathcal{A} \cong_k^{\mathcal{C}} \mathcal{B}$.

Note that this is really a family of three theorems, one for each type of game arising from a resource-indexed arboreal cover $\mathcal{C}$ as in Example 44. Thus in each case, we capture the salient logical equivalences in syntax-free, categorical form.

We return to the general setting. Given a resource-indexed arboreal cover of $\mathcal{E}$ by $\mathcal{C}$, we know by comonadicity that for each k, $\mathcal{C}_k$ is isomorphic to the Eilenberg-Moore category for the comonad G_k. For each $a \in \mathcal{E}$, we can ask if it carries a G_k-coalgebra structure; that is, whether there is a morphism $\alpha \colon a \to G_k a$ satisfying the G_k coalgebra conditions. Moreover, we can ask for the *least* k such that this is the case. We call this the *coalgebra number* of a.

The intuition behind this, as explained in [1, 6], is that the resource parameter is bounding access to the structure, making it more difficult to have a morphism in $\mathcal{E}$ with codomain $G_k a$. So the least k for which this *is* possible is a significant invariant of the structure. This intuition is born out by the following result from [1, 6].

▶ **Theorem 48.**
1. *For the Ehrenfeucht-Fraïssé comonad, the coalgebra number of $\mathcal{A}$ corresponds precisely to the* tree-depth *of $\mathcal{A}$.*
2. *For the pebbling comonad, the coalgebra number of $\mathcal{A}$ corresponds precisely to the* tree-width *of $\mathcal{A}$.*
3. *For the modal comonad, the coalgebra number of $\mathcal{A}$ corresponds precisely to the* modal unfolding depth *of $\mathcal{A}$.*

What underlies these results is the comonadicity of the arboreal covers, which means that the coalgebras are witnesses for the various forms of tree decompositions of structures in $\mathcal{E}$ corresponding to these combinatorial invariants.

As a further illustration of the use of our axiomatic setting, we note that it is possible to give an account of the key notion of *extendability*, used by Rossman in his seminal results on homomorphism preservation [19], at this level of generality. For more details, see [4].

References

1 S. Abramsky, A. Dawar, and P. Wang. The pebbling comonad in finite model theory. In *32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 1–12, 2017.

2 S. Abramsky, R. Jagadeesan, and P. Malacaria. Full abstraction for PCF. *Information and computation*, 163(2):409–470, 2000.

3 S. Abramsky and D. Marsden. Comonadic semantics for guarded fragments. Preprint available at `arXiv:2008.11094`, 2020.

4 S. Abramsky and L. Reggio. Arboreal categories: An axiomatic theory of resources. Extended version. Preprint available at `arXiv:2102.08109`, 2021.

5 S. Abramsky and N. Shah. Relating structure and power: Comonadic semantics for computational resources. In *27th EACSL Annual Conference on Computer Science Logic, CSL, September 4-7, 2018, Birmingham, UK*, pages 2:1–2:17, 2018.

6 S. Abramsky and N. Shah. Relating structure and power: Comonadic semantics for computational resources. Extended version to appear in Journal of Logic and Computation. Preprint available at `arXiv:2010.06496`, 2021.

7 J. Adámek, H. Herrlich, and G.E. Strecker. *Abstract and concrete categories. The joy of cats.* Online edition, 2004.

8 P. Blackburn, M. De Rijke, and Y. Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, 2002.

9 A. Ó Conghaile and A. Dawar. Game comonads & generalised quantifiers. In *29th EACSL Annual Conference on Computer Science Logic, CSL*, volume 183 of *LIPIcs*, pages 16:1–16:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

10 P.J. Freyd and G.M. Kelly. Categories of continuous functors, I. *Journal of Pure and Applied Algebra*, 2(3):169–191, 1972.

11 J.M.E. Hyland and C.-H.L Ong. On full abstraction for PCF: I, II, and III. *Information and computation*, 163(2):285–408, 2000.

12 A. Joyal, M. Nielsen, and G. Winskel. Bisimulation from open maps. *Information and Computation*, 127(2):164–185, 1996.

13 A. Joyal, M. Nielson, and G. Winskel. Bisimulation and open maps. In *Proceedings of 8th Annual IEEE Symposium on Logic in Computer Science*, pages 418–427, 1993.

14 T. Kloks. *Treewidth: computations and approximations*, volume 842 of *Springer Lecture Notes in Computer Science.* Springer Science & Business Media, 1994.

15 L. Libkin. *Elements of finite model theory.* Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2004.

16 S. Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1998.

17 J. Nešetřil and P. Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European Journal of Combinatorics*, 27(6):1022–1041, 2006.

18 E. Riehl. Factorization systems. Notes available at `http://www.math.jhu.edu/~eriehl/factorization.pdf`.

19 B. Rossman. Homomorphism preservation theorems. *J. ACM*, 55(3):15:1–15:53, 2008.

20 ACM SIGLOG. Alonzo Church Award, 2017. URL: `https://siglog.org/winners-of-the-2017-alonzo-church-award/`.