Relating Structure and Power: Comonadic Semantics for Computational Resources Extended Version

Samson Abramsky and Nihil Shah
Oxford University Department of Computer Science

October 15, 2020

Abstract

Combinatorial games are widely used in finite model theory, constraint satisfaction, modal logic and concurrency theory to characterize logical equivalences between structures. In particular, Ehrenfeucht-Fraïssé games, pebble games, and bisimulation games play a central role. We show how each of these types of games can be described in terms of an indexed family of comonads on the category of relational structures and homomorphisms. The index k is a resource parameter which bounds the degree of access to the underlying structure. The coKleisli categories for these comonads can be used to give syntax-free characterizations of a wide range of important logical equivalences. Moreover, the coalgebras for these indexed comonads can be used to characterize key combinatorial parameters: tree-depth for the Ehrenfeucht-Fraïssé comonad, tree-width for the pebbling comonad, and synchronization-tree depth for the modal unfolding comonad. These results pave the way for systematic connections between two major branches of the field of logic in computer science which hitherto have been almost disjoint: categorical semantics, and finite and algorithmic model theory.

1 Introduction

There is a remarkable divide in the field of logic in Computer Science, between two distinct strands: one focusing on semantics and compositionality ("Structure"), the other on expressiveness and complexity ("Power"). It is remarkable because these two fundamental aspects of our field are studied using almost disjoint technical languages and methods, by almost disjoint research communities. We believe that bridging this divide is a major issue in Computer Science, and may hold the key to fundamental advances in the field.

In this paper, we develop a novel approach to relating categorical semantics, which exemplifies the first strand, to finite model theory, which exemplifies the second. It builds on the ideas introduced in [3], but goes much further, showing clearly that there is a strong and robust connection, which can serve as a basis for many further developments. These developments were first introduced in the conference version of this paper [5]. In this journal version, we have added more detailed proofs, and three additional sections demonstrating how game comonads generalize constructions in finite model theory. We have also introduced a new and more satisfactory unified account of back-and-forth equivalences, in terms of open maps.

Acknowledgements

The setting

Relational structures and the homomorphisms between them play a fundamental rôle in finite model theory, constraint satisfaction and database theory. The existence of a homomorphism $A \to B$ is an equivalent formulation of constraint satisfaction, and also equivalent to the preservation of existential positive sentences [12]. This setting also generalizes what has become a central perspective in graph theory [24].

Model theory and deception

In a sense, the purpose of model theory is "deception". It allows us to see structures not "as they really are", *i.e.* up to isomorphism, but only up to *definable properties*, where definability is relative to a logical language $\mathcal{L}$. The key notion is *logical equivalence* $\equiv^{\mathcal{L}}$. Given structures $\mathscr{A}$, $\mathscr{B}$ over the same vocabulary:

$$\mathscr{A} \equiv^{\mathcal{L}} \mathscr{B} \overset{\Delta}{\iff} \forall \varphi \in \mathcal{L}. \ \mathscr{A} \models \varphi \iff \mathscr{B} \models \varphi.$$

If a class of structures $\mathcal{K}$ is definable in $\mathcal{L}$, then it must be saturated under $\equiv^{\mathcal{L}}$. Moreover, for a wide class of cases of interest in finite model theory, the converse holds [30].

The idea of syntax-independent characterizations of logical equivalence is quite a classical one in model theory, exemplified by the Keisler-Shelah theorem [42]. It acquires additional significance in finite model theory, where model comparison games such as Ehrenfeucht-Fraïssé games, pebble games and bisimulation games play a central role [31].

We offer a new perspective on these ideas. We shall study these games, not as external artefacts, but as semantic constructions in their own right. Each model-theoretic comparison game encodes "deception" in terms of limited access to the structure. These limitations are indexed by a parameter which quantifies the resources which control this access. For Ehrenfeucht-Fraïssé games and bisimulation games, this is the number of rounds; for pebble games, the number of pebbles.

Main Results

We now give a conceptual overview of our main results. Technical details will be provided in the following sections.

We shall consider three forms of model comparison game: Ehrenfeucht-Fraïssé games, pebble games and bisimulation games [31]. For each of these notions of game G, and value of the resource parameter k, we shall define a corresponding *comonad* $\mathbb{C}_k$ on the category of relational structures and homomorphisms over some relational vocabulary. For each structure $\mathscr{A}$, $\mathbb{C}_k\mathscr{A}$ is another structure over the same vocabulary, which encodes the limited access to $\mathscr{A}$ afforded by playing the game on $\mathscr{A}$ with k resources. There is always an associated homomorphism $\varepsilon_{\mathscr{A}} : \mathbb{C}_k\mathscr{A} \to \mathscr{A}$ (the *counit* of the comonad), so that $\mathbb{C}_k\mathscr{A}$ "covers" $\mathscr{A}$. Moreover, given a homomorphism $h : \mathbb{C}_k\mathscr{A} \to \mathscr{B}$, there is a *Kleisli coextension* homomorphism $h^* : \mathbb{C}_k\mathscr{A} \to \mathbb{C}_k\mathscr{B}$. This allows us to form the *coKleisli category* $\mathsf{Kl}(\mathbb{C}_k)$ for the comonad. The objects are relational structures, while the morphisms from $\mathscr{A}$ to $\mathscr{B}$ in $\mathsf{Kl}(\mathbb{C}_k)$ are exactly the homomorphisms of the form $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$. Composition of these morphisms uses the Kleisli coextension. The connection between this construction and the corresponding form of game G is expressed by the following result:

Theorem 1.1. *The following are equivalent:*

1. *There is a coKleisli morphism $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$*

2. *Duplicator has a winning strategy for the existential G-game with k resources, played from $\mathscr{A}$ to $\mathscr{B}$.*

The existential form of the game has only a "forth" aspect, without the "back". This means that Spoiler can only play in $\mathscr{A}$, while Duplicator only plays in $\mathscr{B}$. This corresponds to the asymmetric form of the coKleisli morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$. Intuitively, Spoiler plays in $\mathbb{C}_k\mathscr{A}$, which gives them limited access to $\mathscr{A}$, while Duplicator plays in $\mathscr{B}$. The Kleisli coextension guarantees that Duplicator's strategies can always be lifted to $\mathbb{C}_k\mathscr{B}$; while we can always compose a strategy $\mathbb{C}_k\mathscr{A} \to \mathbb{C}_k\mathscr{B}$ with the counit on $\mathscr{B}$ to obtain a coKleisli morphism.

This asymmetric form may seem to limit the scope of this approach, but in fact this is not the case. For each of these comonads $\mathbb{C}_k$, we have the following equivalences:

- $\mathscr{A} \rightleftarrows_k \mathscr{B}$ iff there are coKleisli morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$ and $\mathbb{C}_k\mathscr{B} \to \mathscr{A}$. Note that there need be no relationship between these morphisms.

- $\mathscr{A} \cong_{\mathsf{Kl}(\mathbb{C}_k)} \mathscr{B}$ iff $\mathscr{A}$ and $\mathscr{B}$ are isomorphic in the coKleisli category $\mathsf{Kl}(\mathbb{C}_k)$. This means that there are morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$ and $\mathbb{C}_k\mathscr{B} \to \mathscr{A}$ which are inverses of each other in $\mathsf{Kl}(\mathbb{C}_k)$.

Clearly, $\cong_{\mathsf{Kl}(\mathbb{C}_k)}$ strictly implies $\rightleftarrows_k$. We can also define an intermediate "back-and-forth" equivalence $\leftrightarrow_k$, in terms of the existence of a suitable kind of span of coKleisli morphisms.

For each of our three types of game, there are corresponding fragments $\mathcal{L}_k$ of first-order logic:

- 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 over relational vocabularies with symbols of arity at most 2, $\mathcal{L}_k$ is the modal fragment [9] 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 [31].

We can now state our first main result, in a suitably generic form.

Theorem 1.2. *For finite structures $\mathscr{A}$ and $\mathscr{B}$:*
$$(1) \quad \mathscr{A} \equiv^{\exists\mathcal{L}_k} \mathscr{B} \quad \Longleftrightarrow \quad \mathscr{A} \rightleftarrows_k \mathscr{B}.$$
$$(2) \quad \mathscr{A} \equiv^{\mathcal{L}_k} \mathscr{B} \quad \Longleftrightarrow \quad \mathscr{A} \leftrightarrow_k \mathscr{B}.$$
$$(3) \quad \mathscr{A} \equiv^{\mathcal{L}_k(\#)} \mathscr{B} \quad \Longleftrightarrow \quad \mathscr{A} \cong_{\mathsf{Kl}(\mathbb{C}_k)} \mathscr{B}.$$

Note that this is really a family of three theorems, one for each type of game G. Thus in each case, we capture the salient logical equivalences in syntax-free, categorical form.

We now turn to the significance of indexing by the resource parameter k. When $k \leq l$, we have a natural inclusion morphism $\mathbb{C}_k\mathscr{A} \to \mathbb{C}_l\mathscr{A}$, since playing with k resources is a special case of playing with $l \geq k$ resources. This tells us that the smaller k is, the easier it is to find a morphism $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$. Intuitively, the more we restrict Spoiler's abilities to access the structure of $\mathscr{A}$, the easier it is for Duplicator to win the game.

The contrary analysis applies to morphisms $\mathscr{A} \to \mathbb{C}_k\mathscr{B}$. The smaller k is, the *harder* it is find such a morphism. Note, however, that if $\mathscr{A}$ is a finite structure of cardinality k, then $\mathscr{A} \rightleftarrows_k \mathbb{C}_k\mathscr{A}$. In this case, with k resources we can access the whole of $\mathscr{A}$. What can we say when k is strictly smaller than the cardinality of A?

It turns out that there is a beautiful connection between these indexed comonads and combinatorial invariants of structures. This is mediated by the notion of *coalgebra*, another fundamental (and completely general) aspect of comonads. A coalgebra for a comonad $\mathbb{C}_k$ on a structure $\mathscr{A}$ is a morphism $\mathscr{A} \to \mathbb{C}_k\mathscr{A}$ satisfying certain properties. We define the *coalgebra number* of a structure $\mathscr{A}$, with respect to the indexed family of comonads $\mathbb{C}_k$, to be the least k such that there is a $\mathbb{C}_k$-coalgebra on $\mathscr{A}$.

We now come to our second main result.

Theorem 1.3. *1. For the Ehrenfeucht-Fraïssé comonad, the coalgebra number of $\mathscr{A}$ corresponds precisely to the* tree-depth *of $\mathscr{A}$ [37].*

   2. For the pebbling comonad, the coalgebra number of $\mathscr{A}$ corresponds precisely to the tree-width *of $\mathscr{A}$.*

   3. For the modal comonad, the coalgebra number of $\mathscr{A}$ corresponds precisely to the modal unfolding depth *of $\mathscr{A}$.*

The main idea behind these results is that coalgebras on $\mathscr{A}$ are in bijective correspondence with decompositions of $\mathscr{A}$ of the appropriate form. We thus obtain categorical characterizations of these key combinatorial parameters.

These two results appeared in the first version of this paper [5], and unify logical resources with combinatorial resources. In this expanded version of the paper, we provide two new families of results that can be seen as further aspects of this unification. The first result utilizes a general fact about comonads to link the logical resource characterized by coKleisli morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$ to the combinatorial resource characterized by the coalgebas $\mathscr{A} \to \mathbb{C}_k\mathscr{A}$. For each game comonad $\mathbb{C}_k$, we can define a preorder $\to_k^{\mathbb{C}}$ over the category of structures, where $\mathscr{A} \to_k^{\mathbb{C}} \mathscr{B}$ iff for all structures $\mathscr{C}$ with combinatorial resource parameter $\leq k$, $\mathscr{C} \to \mathscr{A} \Rightarrow \mathscr{C} \to \mathscr{B}$. This preorder, when instantiated to the case of Ehrenfeucht-Fraïssé games and tree-depth, is used extensively in [40]. We demonstrate that this preorder is characterized by the coKleisli category of $\mathbb{C}_k$.

Proposition 1.4. $\mathscr{A} \to_k^{\mathbb{C}} \mathscr{B} \iff \mathbb{C}_k\mathscr{A} \to \mathscr{B}$.

For the second result, we show how coalgebras can be interpreted syntactically. We show that a coalgebra $\mathscr{A} \to \mathbb{C}_k\mathscr{A}$ encodes a witness to the fact that the canonical conjunctive query $Q^{\mathscr{A}}$ of $\mathscr{A}$ is equivalent to a sentence in the logical fragment corresponding to the comonad.

Theorem 1.5. *$\mathscr{A}$ has a $\mathbb{C}_k$-coalgebra iff $Q^{\mathscr{A}}$ can be rewritten into a formula φ, where φ has logical resource at most k.*

The formula φ can be read out from the coalgebra map.

We also show how to rephrase the coalgebra results of Theorem 1.3 in the language of adjunctions.

Theorem 1.6. *Each game comonad arises from an adjunction between the category of structures and a category of tree-ordered structures:*

- *For the Ehrenfeucht-Fraïssé comonad, the associated category is k-height tree-ordered structures.*

- *For the pebbling comonad, the associated category is k-pebble tree-ordered structures.*

- *For the modal comonad, the associated category is k-height tree-ordered Kripke structures.*

In each case, the left adjoint is the evident forgetful functor which forgets the order. The right adjoint is given by the functor part of the comonad.

We have also put the notion of I-morphism introduced in [3], and used extensively in [5], on a more systematic basis, in terms of *relative comonads* [8].

Finally, an important contribution of this expanded version of the paper is to give a unified, general treatment of back-and-forth equivalences, which play a central role in finite model theory. To do this, we use a refined version of the well-known notion of open map bisimulation [27], in which the maps in the span witnessing a bisimulation are required not only to be open, but to be *pathwise embeddings*. This supports a general notion of *property-preserving bisimulation*, which specialises to yield all the back-and-forth equivalences we study here, as well as many other examples, e.g. *guarded bisimulation* [3].

2 Background

We shall now review some background on logics, model comparison games, and comonads.

2.1 Notational preliminaries

A relational vocabulary σ is a set of relation symbols R, each with a specified positive integer arity. A σ-structure $\mathscr{A}$ is given by a set A, the universe of the structure, and for each R in σ with arity n, a relation $R^{\mathscr{A}} \subseteq A^n$. A homomorphism $h : \mathscr{A} \to \mathscr{B}$ is a function $h : A \to B$ such that, for each relation symbol R of arity n in σ, for all $a_1, \ldots, a_n$ in A: $R^{\mathscr{A}}(a_1, \ldots, a_n) \Rightarrow R^{\mathscr{B}}(h(a_1), \ldots, h(a_n))$. We write $\mathcal{R}(\sigma)$ for the category of σ-structures and homomorphisms.

A homomorphism $h : \mathscr{A} \to \mathscr{B}$ is *strong* if for n-ary R in σ and $a_1, \ldots, a_n \in A$, $R^{\mathscr{B}}(h(a_1), \ldots, h(a_n)) \Rightarrow R^{\mathscr{A}}(a_1, \ldots, a_n)$. An *embedding* of relational structures is an injective strong homomorphism. It is easily verified that the embeddings in $\mathcal{R}(\sigma)$ are exactly the extremal monomorphisms; those

monomorphisms m such that, whenever $m = f \circ e$ for an epimorphism e, e is an isomorphism.

A *partial homomorphism* from $\mathscr{A}$ to $\mathscr{B}$ is a finite relation $r \subseteq A \times B$ which is single-valued on its domain $A_0 := \pi_1(r)$, *i.e.* $(a, b), (a, b') \in r \Rightarrow b = b'$, and which is a homomorphism on the substructure of $\mathscr{A}$ determined by A_0. A *partial isomorphism* from $\mathscr{A}$ to $\mathscr{B}$ is a partial homomorphism r such that the converse relation $r^c \subseteq B \times A$ is a partial homomorphism from $\mathscr{B}$ to $\mathscr{A}$.

We shall write $A^{\leq k}$ for the set of non-empty sequences of length $\leq k$ on a set A. We use list notation $[a_1, \ldots, a_j]$ for such sequences, and indicate concatenation by juxtaposition, i.e. ss' is the concatenation of the sequences s and s'. We write $s \sqsubseteq t$ for the prefix ordering on sequences. If $s \sqsubseteq t$, there is a unique s' such that $ss' = t$, which we refer to as the suffix of s in t. For each positive integer n, we define $\mathbf{n} := \{1, \ldots, n\}$.

We shall need a few notions on posets. The comparability relation on a poset $(P, \leq)$ is $x \uparrow y$ iff $x \leq y$ or $y \leq x$. A chain in a poset $(P, \leq)$ is a subset $C \subseteq P$ such that, for all $x, y \in C$, $x \uparrow y$. A *forest* is a poset $(F, \leq)$ such that, for all $x \in F$, the set of predecessors $\downarrow(x) := \{y \in F \mid y \leq x\}$ is a finite chain. The height $\mathsf{ht}(F)$ of a forest F is $\sup_C |C|$, where C ranges over chains in F. A *tree* is a forest with a least element $\bot$ (the root). We write the covering relation for a poset as $\prec$; thus $x \prec y$ iff $x \leq y$, $x \neq y$, and for all z, $x \leq z \leq y$ implies $z = x$ or $z = y$.

2.2 Logic fragments

We shall be considering logics $\mathcal{L}$ which arise as fragments of $\mathcal{L}_{\infty, \omega}$, the extension of first-order logic with infinitary conjunctions and disjunctions, but where formulas contain only finitely many variables. In particular, we will consider the fragments $\mathcal{L}_k$, of formulas with quantifier rank $\leq k$, and $\mathcal{L}^k$, the k-variable fragment. These play a fundamental rôle in finite model theory.

We shall also consider two variants for each of these fragments $\mathcal{L}$. One is the existential positive fragment $\exists \mathcal{L}$, which contains only those formulas of $\mathcal{L}$ built using existential quantifiers, conjunction and disjunction. The other is $\mathcal{L}(\#)$, the extension of $\mathcal{L}$ with counting quantifiers. These have the form $\exists_{\leq n}$, $\exists_{\geq n}$, where the semantics of $\mathscr{A} \models \exists_{\geq n} x. \psi$ is that there exist at least n distinct elements of A satisying ψ.

For modal logic, we will consider $\mathfrak{M}_k$, the *modal fragment* of modal depth $\leq k$. This arises from the standard translation of (multi)modal logic into $\mathcal{L}_{\infty, \omega}$ [10]. Let us fix a relational vocabulary σ with symbols of arity ≤ 2. For each unary symbol P, there will be a corresponding propositional variable p. Formulas are built from these propositional variables by propositional connectives, and modalities $\Box_\alpha$, $\Diamond_\alpha$ corresponding to relations R_α indexed by the binary relation symbols from the given alphabet. Modal formulas φ then admit a translation into formulas $[\![\varphi]\!] = \psi(x)$ in one free variable.

The translation sends propositional variables p to $P(x)$, commutes with the propositional connectives, and sends $\Diamond_\alpha \varphi$ to $\exists y.\, R_\alpha(x,y) \wedge \psi(y)$, and $\Box_\alpha \varphi$ to $\forall y.\, R_\alpha(x,y) \rightarrow \psi(y)$, where $\psi(x) = [\![\varphi]\!]$. This translation is semantics-preserving: given a σ-structure $\mathscr{A}$ and $a \in A$, then $\mathscr{A}, a \models \varphi$ in the sense of Kripke semantics iff $\mathscr{A} \models \psi(a)$ in the standard model-theoretic sense, where $\psi(x) = [\![\varphi]\!]$.

We define the modal depth of a modal formula φ as the maximum nesting depth of modalities occurring in φ. $\mathfrak{M}_k$ is then the image of the translation of modal formulas of modal depth $\leq k$. The existential positive fragment $\exists \mathfrak{M}_k$ arises from the modal sublanguage in which formulas are built from propositional variables using only conjunction, disjunction and the diamond modalities $\Diamond_\alpha$.

Extensions of the modal language with counting capabilities have been studied in the form of *graded modalities* [17]. These have the form $\Diamond_\alpha^n$, $\Box_\alpha^n$, where $\mathscr{A}, a \models \Diamond_\alpha^n \varphi$ if there are at least n R_α-successors of a which satisfy φ. We define $\mathfrak{M}_k(\#)$ to be the extension of the modal fragment with graded modalities.

2.3   Model comparison games

We review the model comparison games we will be concerned with in this paper. We are given structures $\mathscr{A}$ and $\mathscr{B}$. Each game is a 2-person game, played between Spoiler, who is trying show that the structures are different, and Duplicator, who is trying to show they are the same. Each game is played in a number of rounds:

- In the Ehrenfeucht-Fraïssé game, in the i'th round Spoiler chooses an element in one of the structures, and Duplicator then chooses an element in the other structure. Thus after k rounds have been played, we have sequences $[a_1, \ldots, a_k]$ and $[b_1, \ldots b_k]$ of elements from $\mathscr{A}$ and $\mathscr{B}$ respectively. Duplicator wins this play if $r := \{(a_i, b_i) \mid 1 \leq i \leq k\}$ is a partial isomorphism. Duplicator wins the k-round game if they have a winning strategy in the usual sense. The resource parameter here is k, the number of rounds.

- In the pebble game, each player has k pebbles available, which can be "placed" on elements of either structure. We label the pebbles by integers in $\mathbf{k}$. In the i'th round, Spoiler places some pebble p_i on an element of one of the structures. Note that if p_i was previously placed on some other element, the effect is to remove it from that element, and place it on the newly chosen element. Duplicator then places their corresponding pebble p_i on an element of the other structure. Thus the current positions of the pebbles determine "windows" of size bounded by k onto the structures. These windows can slide around over different parts of the structures as moves are played.

After n rounds of the game have been played, we have sequences $[(p_1, a_1), \ldots, (p_n, a_n)]$ and $[(p_1, b_1), \ldots (p_n, b_n)]$ which record the placings of pebbles on elements during the play. The *current placing* of pebble p is the last element in the sequence with first component p. Duplicator wins this play if the relation r determined by the current placings of the pebbles is a partial isomorphism. Duplicator wins the k-pebble game if they have a strategy which is winning after n rounds, for all n. The resource parameter here is k, the number of pebbles.

- Finally, in the bisimulation game, we have structures over a relational signature with symbols of arity at most 2. The game is played between "pointed structures" $(\mathscr{A}, a)$, $(\mathscr{B}, b)$, with specified elements $a \in A$ and $b \in B$. We start with elements (a_0, b_0), where $a_0 = a$ and $b_0 = b$. At each round $i+1$, where the current elements are (a_i, b_i), Spoiler chooses one of the structures, e.g. $\mathscr{A}$, one of the binary relations R_i, and an element a_{i+1} such that $R_i^{\mathscr{A}}(a_i, a_{i+1})$. Duplicator must respond in the other structure, in our example in $\mathscr{B}$ with b_{i+1}, such that $R_i^{\mathscr{B}}(b_i, b_{i+1})$. If Duplicator has no such response available, they lose. Duplicator wins after k rounds if, for all unary predicates P in the signature, we have $P^{\mathscr{A}}(a_i) \iff P^{\mathscr{B}}(b_i)$ for all i with $0 \leq i \leq k$. The resource parameter here is the number of rounds k.

There are classical theorems (Ehrenfeucht-Fraïssé [20, 19], Kolaitis-Vardi [30], Hennessy-Milner [26]) which characterize various logical equivalences in terms of these games. These can be summarized as follows:

Theorem 2.1. *Let $\mathscr{A}$ and $\mathscr{B}$ be structures.*

1. *Duplicator has a winning strategy in the k-round Ehrenfeucht-Fraïssé game from $\mathscr{A}$ to $\mathscr{B}$ iff $\mathscr{A}$ and $\mathscr{B}$ satisfy the same sentences of quantifier rank $\leq k$.*

2. *Duplicator has a winning strategy in the k-pebble game from $\mathscr{A}$ to $\mathscr{B}$ iff $\mathscr{A}$ and $\mathscr{B}$ satisfy the same sentences of k-variable logic.*

3. *Duplicator has a winning strategy in the k-round bisimulation game from $(\mathscr{A}, a)$ to $(\mathscr{B}, b)$ iff $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$ satisfy the same modal formulas of depth $\leq k$.*

We refer to [31, 18, 10] for textbook accounts and extensive bibliographies.

The above games are symmetric or "back-and-forth", since Spoiler can choose which structure to play in at each round. There are also asymmetric (forth only) or *existential* versions, in which Spoiler always plays in $\mathscr{A}$, and Duplicator always responds in $\mathscr{B}$. The winning conditions are modified for these existential versions:

- For the existential Ehrenfeucht-Fraïssé game, the winning condition is that r is a partial homomorphism.

- Similarly for the existential k-pebble game.

- For the k-round *simulation game*, as the existential version of the bisimulation game is called, the winning condition is that, for all unary predicates P, $P(a_i) \Rightarrow P(b_i)$ for all i with $0 \leq i \leq k$.

For these existential games, we obtain a corresponding version of Theorem 2.1.

Theorem 2.2. *Let $\mathscr{A}$ and $\mathscr{B}$ be structures.*

1. *Duplicator has a winning strategy in the existential k-round Ehrenfeucht-Fraïssé game from $\mathscr{A}$ to $\mathscr{B}$ iff for every existential-positive sentence φ of quantifier rank $\leq k$, $\mathscr{A} \models \varphi$ implies $\mathscr{B} \models \varphi$.*

2. *Duplicator has a winning strategy in the existential k-pebble game from $\mathscr{A}$ to $\mathscr{B}$ iff for every existential-positive sentence φ of k-variable logic, $\mathscr{A} \models \varphi$ implies $\mathscr{B} \models \varphi$.*

3. *Duplicator has a winning strategy in the k-round simulation game from $(\mathscr{A}, a)$ to $(\mathscr{B}, b)$ iff for every existential-positive sentence φ of the modal fragment with modal depth $\leq k$, $\mathscr{A} \models \varphi$ implies $\mathscr{B} \models \varphi$.*

2.4 Comonads

We recall that a comonad (G, ε, δ) on a category $\mathcal{C}$ is given by a functor $G : \mathcal{C} \to \mathcal{C}$, and natural transformations $\varepsilon : G \Rightarrow I$ (the counit), and $\delta : G \Rightarrow G^2$ (the comultiplication), subject to the conditions that the following diagrams commute, for all objects A of $\mathcal{C}$:

$$
\begin{array}{ccc}
GA & \xrightarrow{\delta_A} & GGA \\
{\scriptstyle \delta_A}\downarrow & & \downarrow{\scriptstyle G\delta_A} \\
GGA & \xrightarrow{\delta_{GA}} & GGGA
\end{array}
\qquad\qquad
\begin{array}{ccc}
GA & \xrightarrow{\delta_A} & GGA \\
{\scriptstyle \delta_A}\downarrow & & \downarrow{\scriptstyle G\varepsilon_{\mathscr{A}}} \\
GGA & \xrightarrow{\varepsilon_{GA}} & GA
\end{array}
$$

An equivalent formulation is *comonad in Kleisli form* [33]. This is given by an object map G, arrows $\varepsilon_A : GA \to A$ for every object A of $\mathcal{C}$, and a Kleisli coextension operation which takes $f : GA \to B$ to $f^* : GA \to GB$. These must satisfy the following equations:

$$
\varepsilon_A^* = \mathsf{id}_{GA}, \qquad \varepsilon \circ f^* = f, \qquad (g \circ f^*)^* = g^* \circ f^*.
$$

We can then extend G to a functor by $Gf = (f \circ \varepsilon)^*$; and if we define the comultiplication $\delta : G \Rightarrow G^2$ by $\delta_A = \mathsf{id}_{GA}^*$, then (G, ε, δ) is a comonad in the standard sense. Conversely, given a comonad (G, ε, δ), we can define the

10

coextension by $f^* = Gf \circ \delta_A$. This allows us to define the coKleisli category $\mathsf{Kl}(G)$, with objects the same as those of $\mathcal{C}$, and morphisms from A to B given by the morphisms in $\mathcal{C}$ of the form $GA \to B$. Kleisli composition of $f : GA \to B$ with $g : GB \to C$ is given by $g \bullet f := g \circ f^*$.

3 Game Comonads

We shall now define the three comonads we shall study, corresponding to the three forms of model comparison game described in the previous section.

3.1 The Ehrenfeucht-Fraïssé Comonad

We shall define a comonad $\mathbb{E}_k$ on $\mathcal{R}(\sigma)$ for each positive integer k. It will be convenient to define $\mathbb{E}_k$ in Kleisli form. For each structure $\mathscr{A}$, we define a new structure $\mathbb{E}_k\mathscr{A}$, with universe $\mathbb{E}_k A := A^{\leq k}$. We define the map $\varepsilon_\mathscr{A} : \mathbb{E}_k A \to A$ by $\varepsilon_\mathscr{A}[a_1, \ldots, a_j] = a_j$. For each relation symbol R of arity n, we define $R^{\mathbb{E}_k\mathscr{A}}$ to be the set of n-tuples $(s_1, \ldots, s_n)$ of sequences which are pairwise comparable in the prefix ordering, and such that $R^\mathscr{A}(\varepsilon_\mathscr{A}s_1, \ldots, \varepsilon_\mathscr{A}s_n)$. Finally, we define the coextension. Given a homomorphism $f : \mathbb{E}_k\mathscr{A} \to \mathscr{B}$, we define $f^* : A^{\leq k} \to B^{\leq k}$ by $f^*[a_1, \ldots, a_j] = [b_1, \ldots, b_j]$, where $b_i = f[a_1, \ldots, a_i]$, $1 \leq i \leq j$.

Proposition 3.1. *The triple $(\mathbb{E}_k, \varepsilon, (\cdot)^*)$ is a comonad in Kleisli form.*

Proof. It is well known that non-empty lists form a comonad on **Set** [7]. This is easily adapted to showing that non-empty lists of length $\leq k$ form a comonad. To lift this to $\mathcal{R}(\sigma)$, we must show that each $\varepsilon_\mathscr{A}$ is a homomorphism, and that if $f : \mathbb{E}_k\mathscr{A} \to \mathscr{B}$ is a homomorphism, so is f^*. The first follows directly from the definition of $R^{\mathbb{E}_k\mathscr{A}}$ for each R in σ. For the second, if $R^{\mathbb{E}_k\mathscr{A}}(s_1, \ldots, s_n)$, then $s_i \uparrow s_j$ implies $f^*(s_i) \uparrow f^*(s_j)$, and since f is a homomorphism, $R^{\mathbb{E}_k\mathscr{A}}(s_1, \ldots, s_n)$ implies $R^\mathscr{B}(\varepsilon f^*(s_1), \ldots, \varepsilon f^*(s_n))$, so $R^{\mathbb{E}_k\mathscr{B}}(f^*(s_1), \ldots, f^*(s_n))$ as required. $\square$

Intuitively, an element of $A^{\leq k}$ represents a play in $\mathscr{A}$ of length $\leq k$. A coKleisli morphism $\mathbb{E}_k\mathscr{A} \to \mathscr{B}$ represents a Duplicator strategy for the existential Ehrenfeucht-Fraïssé game with k rounds, where Spoiler plays only in $\mathscr{A}$, and $b_i = f[a_1, \ldots, a_i]$ represents Duplicator's response in $\mathscr{B}$ to the i'th move by Spoiler. The winning condition for Duplicator in this game is that, after k rounds have been played, the *induced relation* $\{(a_i, b_i) \mid 1 \leq i \leq k\}$ is a partial homomorphism from $\mathscr{A}$ to $\mathscr{B}$.

These intuitions are confirmed by the following result.

Theorem 3.2. *The following are equivalent:*

1. *There is a homomorphism $\mathbb{E}_k\mathscr{A} \to \mathscr{B}$.*

2. Duplicator has a winning strategy for the existential Ehrenfeucht-Fraïssé game with k rounds, played from $\mathscr{A}$ to $\mathscr{B}$.

Proof. Suppose Spoiler plays $a_1, \ldots, a_k$, then for each $i \in \{1, \ldots, k\}$, Duplicator responds with $b_i = b_j$ if $a_j = a_i$ for some $j < i$ [1] or $b_i = f[a_1, \ldots, a_i]$ otherwise. We must show that $\gamma = \{(a_i, b_i) \mid 1 \leq i \leq k\}$ is a partial homomorphism. By construction, $b_i = b_j$ if $a_i = a_j$, so γ is a well-defined partial function. Let $R \in \sigma$ and suppose $R^{\mathscr{A}}(a_{i_1}, \ldots, a_{i_n})$ with $i_j \in \{1, \ldots, k\}$ for all $j \leq n$. Let s_{i_j} be the minimal subsequence of $s = [a_1, \ldots, a_k]$ that ends in a_{i_j}. Since each s_{i_j} is a subsequence of s, each pair of s_{i_j} is prefix comparable and by definition, $\epsilon_A(s_{i_j}) = a_{i_j}$, so $R^{\mathbb{E}_k \mathscr{A}}(s_{i_1}, \ldots, s_{i_n})$. By f being a homomorphism, $R^{\mathscr{B}}(b_{i_1}, \ldots, b_{i_n})$, so γ is a partial homomorphism.

Conversely, suppose Duplicator has a winning strategy in the n-round EF-game (with $n \leq k$). For every possible set of Spoiler moves in the n-round game, i.e. for every sequence $s = [a_1, \ldots, a_n]$, there exists a sequence $t = [b_1, \ldots, b_n]$ such that $\gamma = \{(a_i, b_i) \mid 1 \leq i \leq n\}$ is a partial homomorphism. Let γ_s be the partial homomorphism resulting from Duplicator's responses to Spoiler's play s. We define $f : \mathbb{E}_k \mathscr{A} \to \mathscr{B}$ by $f(s_i) = \gamma_s(a_i)$ where $s_i \sqsubseteq s$ ending in a_i. Since γ_s is a partial homomorphism for all $s \in \mathbb{E}_k \mathscr{A}$, f is a homomorphism. Namely, suppose $s_1, \ldots, s_m$ are such that $R^{\mathbb{E}_k \mathscr{A}}(s_1, \ldots, s_m)$. By the pairwise comparability condition, there exists some $\sqsubseteq$-greatest $s = [a_1, \ldots, a_l]$ with $l \leq n$ amongst the s_i. Since $s_i \sqsubseteq s$, the last move of s_i, $\epsilon_{\mathscr{A}}(s_i)$ must be in $\{a_1, \ldots, a_l\}$. By the compatiblity condition, we have that $R^{\mathscr{A}}(\epsilon_A(s_1), \ldots, \epsilon_A(s_m))$. Hence, since γ_s is a partial homomorphism, $R^{\mathscr{B}}(\gamma_s(\epsilon_{\mathscr{A}}(s_1)), \ldots, \gamma_s(\epsilon_{\mathscr{A}}(s_m)))$. Since $f(s_i) = \gamma_s(\epsilon_A(s_i))$, we have that $R^{\mathscr{B}}(f(s_1), \ldots, f(s_m))$. $\qquad \square$

3.2 The Pebbling Comonad

We now turn to the case of pebble games. The following construction appeared in [3]. Given a structure $\mathscr{A}$, we define $\mathbb{P}_k \mathscr{A}$, which will represent plays of the k-pebble game on $\mathscr{A}$.[2] The universe is $(\mathbf{k} \times A)^+$, the set of finite non-empty sequences of moves (p, a), where $p \in \mathbf{k}$ is a pebble index, and $a \in A$. We shall use the notation $s = [(p_1, a_1), \ldots, (p_n, a_n)]$ for these sequences, which may be of arbitrary length. Thus the universe of $\mathbb{P}_k \mathscr{A}$ is always infinite, even if $\mathscr{A}$ is a finite structure. This is unavoidable, by [3, Theorem 7]. We define $\varepsilon_{\mathscr{A}} : \mathbb{P}_k A \to A$ to send a play $[(p_1, a_1), \ldots, (p_n, a_n)]$ to a_n, the A-component of its last move.

Given an n-ary relation $R \in \sigma$, we define $R^{\mathbb{P}_k \mathscr{A}}(s_1, \ldots, s_n)$ iff (1) the s_i are pairwise comparable in the prefix ordering; (2) the pebble index of the

---

[1] We could eliminate this case by imposing an additional *I-morphism* condition on the homomorphisms of type $\mathbb{E}_k \mathscr{A} \to \mathscr{B}$. This condition is discussed in section 4.

[2] In [3] we used the notation $\mathbb{T}_k$ for this comonad.

last move in each s_i does not appear in the suffix of s_i in s_j for any $s_j \sqsupseteq s_i$; and (3) $R^{\mathscr{A}}(\varepsilon_{\mathscr{A}}(s_1), \ldots, \varepsilon_{\mathscr{A}}(s_n))$.

Note that this differs from the definition of $R^{\mathbb{E}_k \mathscr{A}}$ just in condition (2). Intuitively, Spoiler moves by placing one from a fixed set of pebbles on an element of A; Duplicator responds by placing their matching pebble on an element of the other structure. Thus there is a "sliding window" on the structures, of fixed size. It is this size which bounds the resource, not the length of the play. In particular, placing a pebble on an element implies removing it from the element on which it was previously placed. Condition (2) enforces that only the *current* position of a pebble is taken into account when deciding whether elements are related.

Finally, given a homomorphism $f : \mathbb{P}_k \mathscr{A} \to \mathscr{B}$, we define $f^* : \mathbb{P}_k A \to \mathbb{P}_k B$ by
$f^*[(p_1, a_1), \ldots, (p_j, a_j)] = [(p_1, b_1), \ldots, (p_j, b_j)]$, where $b_i = f[(p_1, a_1), \ldots, (p_i, a_i)]$, $1 \leq i \leq j$.

Proposition 3.3. *The triple* $(\mathbb{P}_k, \varepsilon, (\cdot)^*)$ *is a comonad in Kleisli form.*

This is [3, Theorem 4]. The following is [3, Theorem 13].

Theorem 3.4. *The following are equivalent:*

1. *There is a homomorphism* $\mathbb{P}_k \mathscr{A} \to \mathscr{B}$.

2. *There is a winning strategy for Duplicator in the existential k-pebble game from* $\mathscr{A}$ *to* $\mathscr{B}$.

3.3 The Modal Comonad

For the modal case, we assume that the relational vocabulary σ contains only symbols of arity at most 2. We can thus regard a σ-structure as a Kripke structure for a multi-modal logic, where the universe is thought of as a set of worlds, each binary relation symbol R_α gives the accessibility relation for one of the modalities, and each unary relation symbol P give the valuation for a corresponding propositional variable. If there are no unary symbols, such structures are exactly the labelled transition systems widely studied in concurrency [35].

Modal logic localizes its notion of satisfaction in a structure to a world. We shall reflect this by using the category of *pointed relational structures* $\mathcal{R}_\star(\sigma)$. Objects of this category are pairs $(\mathscr{A}, a)$ where $\mathscr{A}$ is a σ-structure and $a \in A$. Morphisms $h : (\mathscr{A}, a) \to (\mathscr{B}, b)$ are homomorphisms $h : \mathscr{A} \to \mathscr{B}$ such that $h(a) = b$. Of course, the same effect could be achieved by expanding the vocabulary σ with a constant, but pointed categories appear in many mathematical contexts.

For each $k \geq 0$, we shall define a comonad $\mathbb{M}_k$, where $\mathbb{M}_k(\mathscr{A}, a)$ corresponds to unravelling the structure $\mathscr{A}$, starting from a, to depth k. The universe of $\mathbb{M}_k(\mathscr{A}, a)$ comprises the unit sequence $[a]$, which is the distinguished

element, together with all sequences of the form $[a_0, \alpha_1, a_1, \ldots, \alpha_j, a_j]$, where $a = a_0$, $1 \leq j \leq k$, and $R^{\mathscr{A}}_{\alpha_i}(a_i, a_{i+1})$, $0 \leq i < j$. Note that the universe of $\mathbb{M}_0(\mathscr{A}, a)$ is $\{[a]\}$.

The map $\varepsilon_{\mathscr{A}} : \mathbb{M}_k(A, a) \to (A, a)$ sends a sequence to its last element. Unary relation symbols P are interpreted by $P^{\mathbb{M}_k(\mathscr{A}, a)}(s)$ iff $P^{\mathscr{A}}(\varepsilon_{\mathscr{A}} s)$. For binary relations R_α, the interpretation is $R^{\mathbb{M}_k(\mathscr{A}, a)}_\alpha(s, t)$ iff for some $a' \in A$, $t = s[\alpha, a']$. Given a morphism $f : \mathbb{M}_k(\mathscr{A}, a) \to (\mathscr{B}, b)$, we define $f^* : \mathbb{M}_k(\mathscr{A}, a) \to \mathbb{M}_k(\mathscr{B}, b)$ recursively by $f^*[a] = [b]$, $f^*(s[\alpha, a']) = f^*(s)[\alpha, b']$ where $b' = f(s[\alpha, a'])$. This is well-defined since f is a morphism by assumption.

Proposition 3.5. *The triple* $(\mathbb{M}_k, \varepsilon, (\cdot)^*)$ *is a comonad in Kleisli form on* $\mathcal{R}_\star(\sigma)$.

We recall the notion of *simulation* between Kripke structures [10]. Given structures $\mathscr{A}$, $\mathscr{B}$, we define relations $\preceq_k \subseteq A \times B$, $k \geq 0$, by induction on k: $a \preceq_{k+1} b$ iff

1. for all unary P, $P^{\mathscr{A}}(a)$ implies $P^{\mathscr{B}}(b)$

2. for all R_α, if $R^{\mathscr{A}}_\alpha(a, a')$, then for some b', $R^{\mathscr{B}}_\alpha(b, b')$ and $a' \preceq_k b'$.

The base case $a \preceq_0 b$ holds whenever the first condition is satisfied. It is standard that these relations are equivalently formulated in terms of a modified existential Ehrenfeucht-Fraïssé game [10, 23].

Theorem 3.6. *Let* $\mathscr{A}$, $\mathscr{B}$ *be Kripke structures, with* $a \in A$ *and* $b \in B$, *and* $k \geq 0$. *The following are equivalent:*

1. *There is a homomorphism* $f : \mathbb{M}_k(\mathscr{A}, a) \to (\mathscr{B}, b)$.

2. $a \preceq_k b$.

3. *There is a winning strategy for Duplicator in the k-round simulation game from* $(\mathscr{A}, a)$ *to* $(\mathscr{B}, b)$.

Proof. $(1) \Rightarrow (2)$ We prove this by induction on k. For the base case $k = 0$, by (1), there exists a homomorphism, $f : \mathbb{M}_0(\mathscr{A}, a) \to (\mathscr{B}, b)$. Hence, for all unary P, $P^{\mathscr{A}}(a) \Rightarrow P^{\mathbb{M}_1(\mathscr{A}, a)}([a]) \Rightarrow P^{\mathscr{B}}(f[a]) \Rightarrow P^{\mathscr{B}}(b)$, so $a \preceq_0 b$. For the inductive step, let $k' = k + 1$ and assume there is a homomorphism $f : \mathbb{M}_{k'}(\mathscr{A}, a) \to (\mathscr{B}, b)$. Suppose that $R^{\mathscr{A}}_\alpha(a, a')$, then there exists a $b' = f[a, \alpha, a']$ such that $R^{\mathscr{B}}_\alpha(b, b')$. Consider the homomorphism $f_{\alpha, a'} : \mathbb{M}_k(\mathscr{A}, a') \to (\mathscr{B}, b')$ given by $f[a'] = b'$ and $f_{\alpha, a'}(s) = f([a, \alpha, a']s)$. Applying the inductive hypothesis, we obtain that $a' \preceq_k b'$, verifying condition 2 for a simulation. To verify condition 1, by the definition of $\mathbb{M}_{k'}(\mathscr{A}, a)$ and f being a homomorphism we have: $P^{\mathscr{A}}(a) \Rightarrow P^{\mathbb{M}_{k'}(\mathscr{A}, a)}([a]) \Rightarrow P^{\mathscr{B}}(f[a]) \Rightarrow P^{\mathscr{B}}(b)$. Therefore, $a \preceq_{k+1} b$ as desired.

$(2) \Rightarrow (3)$ This is a standard result, and is detailed in section 2.7 of [10].

$(3) \Rightarrow (1)$ We construct f_i for $i \leq k$ by induction. For the base case $k = 0$, define $f_0 : \mathbb{M}_0(\mathscr{A}, a) \to (\mathscr{B}, b)$ as $f_0[a] = b$. Suppose, by the inductive hypothesis, $f_k : \mathbb{M}_k(\mathscr{A}, a) \to (\mathscr{B}, b)$ has been constructed. Let $k' = k + 1$ and we define $f_{k'} : \mathbb{M}_{k'}(\mathscr{A}, a) \to (\mathscr{B}, b)$ by splitting into two cases. For the first case, consider $t \in \mathbb{M}_{k'}(\mathscr{A}, a)$ such that $t = s[\alpha, a']$ for some $s \in \mathbb{M}_k(\mathscr{A}, a)$, $R_\alpha \in \sigma$ and $R_\alpha^{\mathscr{A}}(a, a')$. We can take s as the record of the Spoiler's moves in the first k rounds of the simulation game. In this case, we let $f_{k'}(t) = b'$ where b' is Duplicator's response to Spoiler choosing the α-transition to a' in the k'-th round. For the second case, consider $t \in \mathbb{M}_k(\mathscr{A}, a) \subseteq \mathbb{M}_{k'}(\mathscr{A}, a)$. In this case, define $f_{k'}(t) = f_k(t)$. We need to verify $f_{k'}$ is indeed a homomorphism. Suppose $s, t \in \mathbb{M}_{k'}(\mathscr{A}, a)$ are such that $R_\alpha^{\mathbb{M}_{k'}(\mathscr{A}, a)}(s, t)$. If $s, t \in \mathbb{M}_k(\mathscr{A}, a) \subseteq \mathbb{M}_{k'}(\mathscr{A}, a)$, then $f_{k'}(s) = f_k(s)$ and $f_{k'}(t) = f_k(t)$. Therefore, by the inductive hypothesis that f_k is a homomorphism, $R_\alpha^{\mathscr{B}}(f_{k'}(s), f_{k'}(t))$, and moreover, for each unary predicate P, $P^{\mathbb{M}_k(\mathscr{A}, a)}(t)$ implies $P^{(\mathscr{B}, b)}(f_{k'}(t))$. By the definition of $R_\alpha^{\mathbb{M}_{k'}(\mathscr{A}, a)}$, the only other case is $t = s[\alpha, a']$. Since $f_{k'}(t)$ is Duplicator's response to the k'-round in her winning strategy, we can conclude that $R_\alpha^{\mathscr{B}}(f_{k'}(s), f_{k'}(t))$, and for each unary predicate P, $P^{\mathbb{M}_k(\mathscr{A}, a)}(t)$ implies $P^{(\mathscr{B}, b)}(f_{k'}(t))$. $\square$

4 I-morphisms

Before turning to logical equivalences, there is a technical issue which needs to be addressed.

A coKleisli morphism $f : \mathbb{E}_k \mathscr{A} \to \mathscr{B}$ is an I-morphism if $s \sqsubseteq t$ and $\varepsilon_{\mathscr{A}}(s) = \varepsilon_{\mathscr{A}}(t)$ implies that $f(s) = f(t)$.

Proposition 4.1. *Consider $\sigma^+ = \sigma \cup \{I\}$, where I is a binary relation symbol not in σ. If we interpret $I^{\mathscr{A}}$ and $I^{\mathscr{B}}$ as the identity relations on A and B, then f is an I-morphism iff it is a σ^+-homomorphism.*

Proof. Suppose f is an I-morphism, and $I^{\mathbb{E}_k \mathscr{A}}(s, t)$. By definition, this means that $s \uparrow t$ and $I^{\mathscr{A}}(\varepsilon_{\mathscr{A}}(s), \varepsilon_{\mathscr{A}}(t))$, *i.e.* $\varepsilon_{\mathscr{A}}(s) = \varepsilon_{\mathscr{A}}(t)$. By the I-morphism definition, this implies $f(s) = f(t)$, and by the interpretation of $I^{\mathscr{B}}$ as the identity relation, $I^{\mathscr{B}}(f(s), f(t))$. Hence, f preserves I, and is therefore a σ^+-homomorphism. The proof of the converse is similar. $\square$

The significance of the I-morphism condition becomes apparent from the following proposition.

Proposition 4.2. *If $f : \mathbb{E}_k \mathscr{A} \to \mathscr{B}$ and $g : \mathbb{E}_k \mathscr{B} \to \mathscr{A}$ are I-morphisms, then $f^*(s) = t$, $g^*(t) = s$ imply that (s, t) defines a partial isomorphism from $\mathscr{A}$ to $\mathscr{B}$.*

Proof. Let $s = [a_1, \ldots, a_n]$ and $t = [b_1, \ldots, b_n]$ with $n \leq k$, then the partial function defined by (s,t) is $\gamma : a_i \mapsto b_i$. Suppose $a_j = a_l$ for some $j, l \leq n$ and let s_j, s_l be the subsequences of s ending in a_j, a_l (respectively). Since $s_j \uparrow s_l$ and $\epsilon_A(s_j) = a_j = a_l = \epsilon_A(s_l)$, by the I-morphism condition, $f(s_j) = f(s_l)$. By definition of f and γ, $b_j = b_l$. Hence, γ is well-defined. Suppose $R \in \sigma$ is an m-ary relation symbol with $R^{\mathscr{A}}(a_{i_1}, \ldots, a_{i_m})$ for $a_{i_j} \in s$ and let s_{i_j} denote the prefix of s such that $\epsilon_A(s_{i_j}) = a_{i_j}$. Since each s_{i_j} is a prefix of s, the set of all s_{i_j} is pairwise comparable, so $R^{\mathbb{E}_k \mathscr{A}}(s_{i_1}, \ldots, s_{i_m}) \Rightarrow R^{\mathscr{B}}(f(s_{i_1}), \ldots, f(s_{i_m})) \Rightarrow R^{\mathscr{B}}(b_{i_1}, \ldots, b_{i_m})$, since $f(s_{i_j}) = b_{i_j}$ follows from $f^*(s) = t$. Hence, γ preserves relations, and therefore is a partial homomorphism. Similarly, define $\gamma^{-1} : b_i \mapsto a_i$, which is also well-defined and a partial homomorphism. By construction, γ and γ^{-1} are inverses, so (s,t) defines a partial isomorphism from $\mathscr{A}$ to $\mathscr{B}$. $\qquad\square$

In the light of this result, the reader may wonder why we do not simply work with a sub-comonad of $\mathbb{E}_k$, in which only non-repeating sequences are allowed. The problem with this is that the functorial action of $\mathbb{E}_k$ on non-injective maps will produce repeating sequences. The same problem arises with the coKleisli extension.

A similar notion of I-morphism applies to the pebbling comonad [3]. Note that, if $I^{\mathscr{A}}$ is the identity relation on $\mathscr{A}$, $I^{\mathbb{P}_k \mathscr{A}}(s, t)$ holds if $s = s'[(p, a)]$, $t = t'[(p', a)]$, s and t are comparable in the prefix order, say $s \sqsubseteq t$, and the placing of p in s is still current in t. If $p \neq p'$, this means that two different pebbles are placed on the same element of $\mathscr{A}$. We call such a sequence *duplicating*. Duplicating sequences are to $\mathbb{P}_k$ what repeating sequences are to $\mathbb{E}_k$.

Dan Marsden and Tomáš Jakl have shown that it is possible to define a retract of the $\mathbb{E}_k$ comonad in a more complicated fashion, which does eliminate repeating sequences. However, it is not clear how this can be extended to the pebbling comonad.

We shall develop an alternative, general approach, based on *relative comonads* [8]. We shall not in fact need the full generality of relative comonads. We can take advantage of the fact that our comonads are defined uniformly in the relational vocabulary. Given a relational vocabulary σ, there is a full and faithful embedding $J : \mathcal{R}(\sigma) \to \mathcal{R}(\sigma^+)$ such that $I^{J\mathscr{A}}$ is the identity on A. We define the J-relative comonad $\mathbb{E}_k^+$ on $\mathcal{R}(\sigma)$ to have the following coKleisli category: objects are those of $\mathcal{R}(\sigma)$, morphisms from $\mathscr{A}$ to $\mathscr{B}$ are $\mathcal{R}(\sigma)$-morphisms $\mathbb{E}_k^I J\mathscr{A} \to J\mathscr{B}$, where $\mathbb{E}_k^I$ is our $\mathbb{E}_k$ construction applied to $\mathcal{R}(\sigma^+)$. The coKleisli extension and identity are defined in the evident fashion. The morphisms in this coKleisli category are exactly the I-morphisms in the coKleisli category for the comonad $\mathbb{E}_k$ on $\mathcal{R}(\sigma)$.

In general, given a functor $J : \mathcal{C} \to \mathcal{D}$, and a comonad G on $\mathcal{D}$, we can obtain a relative comonad on $\mathcal{C}$ in this fashion [8]. In particular, this construction can be carried out in exactly the same fashion for the pebbling

16

comonad. The coKleisli morphisms in this case are the non-duplicating morphisms.

5 Logical Equivalences I

We now show how our game comonads can be used to give syntax-free characterizations of a range of logical equivalences, which play a central rôle in finite model theory and modal logic.

Two equivalences can be defined uniformly for any indexed family of comonads $\mathbb{C}_k$:

- $\mathscr{A} \rightleftarrows_k^{\mathbb{C}} \mathscr{B}$ iff there are coKleisli morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$ and $\mathbb{C}_k\mathscr{B} \to \mathscr{A}$. Note that there need be no relationship between these morphisms. This is simply the equivalence induced by the preorder collapse of the coKleisli category.

- $\mathscr{A} \cong_k^{\mathbb{C}} \mathscr{B}$ iff $\mathscr{A}$ and $\mathscr{B}$ are isomorphic in the coKleisli category $\mathsf{Kl}(\mathbb{C}_k)$. This means that there are morphisms $\mathbb{C}_k\mathscr{A} \to \mathscr{B}$ and $\mathbb{C}_k\mathscr{B} \to \mathscr{A}$ which are inverses of each other in $\mathsf{Kl}(\mathbb{C}_k)$.

Clearly, $\cong_k^{\mathbb{C}}$ strictly implies $\rightleftarrows_k^{\mathbb{C}}$.

We introduced a number of logic fragments in Section 2.2. Each of these logics $\mathcal{L}$ induces an equivalence on structures in $\mathcal{R}(\sigma)$:

$$\mathscr{A} \equiv^{\mathcal{L}} \mathscr{B} \overset{\Delta}{\iff} \forall \varphi \in \mathcal{L}.\ \mathscr{A} \models \varphi \iff \mathscr{B} \models \varphi.$$

Our aim is to characterize these equivalences in terms of our game comonads, and more specifically, to use morphisms in the coKleisli categories as witnesses for these equivalences.

We shall use the two equivalences $\rightleftarrows_k^{\mathbb{C}}$ and $\cong_k^{\mathbb{C}}$ induced by comonads as described above to characterize the logical equivalences induced by the existential-positive and counting versions of these logic fragments, $\exists \mathcal{L}_k$ and $\mathcal{L}_k(\#)$, respectively. We shall use the relative versions $\mathbb{E}_k^+$ and $\mathbb{P}_k^+$ of the Ehrenfeucht-Fraïssé and pebbling comonads, as described in the previous section. These impose the condition of being I-morphisms, which will be important to get a precise match with the winning conditions of the corresponding logical games.

We shall defer the treatment of back-and-forth equivalences to Section 10, as it will build on material to be developed later.

We now turn to a detailed study of each of our comonads in turn.

5.1 The Ehrenfeucht-Fraïssé comonad

We begin by using the I-morphism condition enforced by $\mathbb{E}_k^+$ to sharpen Theorem 3.2 from an equivalence to a bijective correspondence.

Theorem 5.1. *There is a bijective correspondence between:*

1. $\mathbb{E}_k^+$ *coKleisli morphisms from $\mathscr{A}$ to $\mathscr{B}$.*

2. *Winning strategies for Duplicator for the existential Ehrenfeucht-Fraïssé game with k rounds, played from $\mathscr{A}$ to $\mathscr{B}$.*

Proof. The proof proceeds as for Theorem 3.2, but we now use the fact that the I-morphism condition ensures that if $f^*(s) = t$, the correspondence $s_i \mapsto t_i$ is single-valued, and hence satisfies the partial homomorphism winning condition for the existential game. Conversely, a winning strategy will induce a coKleisli map satisfying the I-morphism condition. $\square$

We now recall the *bijection game* [25]. In this variant of the Ehrenfeucht-Fraïssé game, Spoiler wins if the two structures have different cardinality. Otherwise, at round i, Duplicator chooses a bijection ψ_i between A and B, and Spoiler chooses an element a_i of A. This determines the choice by Duplicator of $b_i = \psi_i(a_i)$. Duplicator wins after k rounds if the relation $\{(a_i, b_i) \mid 1 \leq i \leq k\}$ is a partial isomorphism.

Proposition 5.2. *The following are equivalent, for finite structures $\mathscr{A}$ and $\mathscr{B}$:*

1. $\mathscr{A} \cong_k^{\mathbb{E}} \mathscr{B}$.

2. *There is a winning strategy for Duplicator in the k-round bijection game.*

Proof. Assuming (1), we have I-morphisms $f : \mathbb{E}_k \mathscr{A} \to \mathscr{B}$ and $g : \mathbb{E}_k \mathscr{B} \to \mathscr{A}$ with $g^* = f^{*-1}$. For each $s \in \{[]\} \cup A^{<k}$, where $[]$ is the empty sequence, and $A^{<k}$ is the set of sequences over A of length strictly less than k, we can define a map $\psi_s : A \to B$, by $\psi_s(a) = f(s[a])$. This is a bijection, with inverse defined similarly from g. These bijections provide a strategy for Duplicator. Since each $(s, f^*(s))$ is a partial isomorphism, this is a winning strategy.

Conversely, a winning strategy provides bijections ψ_s, which we can use to define f by $f(s[a]) = \psi_s(a)$. The winning conditions imply that this is an I-isomorphism in the coKleisli category. $\square$

We can now state our main result on logical equivalences for the Ehrenfeucht-Fraïssé comonad.

Theorem 5.3.   1. *For all structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\exists \mathcal{L}_k} \mathscr{B} \iff \mathscr{A} \rightleftarrows_k^{\mathbb{E}} \mathscr{B}$.*

2. *For all finite structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\mathcal{L}_k(\#)} \mathscr{B} \iff \mathscr{A} \cong_k^{\mathbb{E}} \mathscr{B}$.*

Proof. (1) In [29], it is shown that $\mathscr{A} \equiv^{\exists \mathcal{L}_k} \mathscr{B}$ iff Duplicator has a winning strategy in the k-round existential Ehrenfeucht-Fraïssé game from $\mathscr{A}$ to $\mathscr{B}$. Combining this with Theorem 3.2 yields the result.

(2) In [25] (see also the exposition in [31][Theorem 8.13]) it is shown that $\mathscr{A} \equiv^{\mathcal{L}_k(\#)} \mathscr{B}$ iff Duplicator has a winning strategy in the k-round bijection game. Combining this with Proposition 5.2 yields the result. $\qquad\square$

5.2 The Pebbling Comonad

We now state the following result, characterizing the equivalences induced by finite-variable logics $\mathcal{L}^k$.

Theorem 5.4. *1. For all structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\exists \mathcal{L}^k} \mathscr{B} \iff \mathscr{A} \rightleftarrows_k^{\mathbb{P}} \mathscr{B}$.*

2. For all finite structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\mathcal{L}^k(\#)} \mathscr{B} \iff \mathscr{A} \cong_k^{\mathbb{P}} \mathscr{B}$.

Proof. This follows from Theorems 14 and 18 of [3]. $\qquad\square$

5.3 The Modal Comonad

Corresponding to the graded modal logic described in Section 2.2, a notion of *graded bisimulation* is given in [17]. This is in turn related to *resource bisimulation* [14], which has been introduced in the concurrency setting. The two notions are shown to coincide for image-finite Kripke structures in [6], where it is also shown that they can be presented in a simplified form. We recall that a Kripke structure $\mathscr{A}$ is image-finite if for all $a \in A$ and R_α, $R_\alpha^{\mathscr{A}}(a) := \{a' \mid R^{\mathscr{A}}(a, a')\}$ is finite.

Adapting the results in [6], we define approximants $\sim_k^{\mathbf{g}}$ for graded bisimulation, by induction on $k \geq 0$. We define $a \sim_{k+1}^{\mathbf{g}} b$ to hold iff:

1. for all P, $P^{\mathscr{A}}(a)$ iff $P^{\mathscr{B}}(b)$

2. for all R_α, there is a bijection $\theta : R_\alpha^{\mathscr{A}}(a) \cong R_\alpha^{\mathscr{B}}(b)$ such that, for all $a' \in R_\alpha^{\mathscr{A}}(a)$, $a' \sim_k^{\mathbf{g}} \theta(a')$.

The base case, $a \sim_0^{\mathbf{g}} b$, holds whenever only the first condition is satisfied. We can also define a corresponding graded bisimulation game between $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$.

- At round 0, the elements $a_0 = a$ and $b_0 = b$ are chosen. Duplicator wins if for all P, $P^{\mathscr{A}}(a)$ iff $P^{\mathscr{B}}(b)$, otherwise Spoiler wins.

- At round $i + 1$, Spoiler chooses some R_α, and Duplicator chooses a bijection $\theta_i : R_\alpha^{\mathscr{A}}(a_i) \cong R_\alpha^{\mathscr{B}}(b_i)$. If there is no such bijection, Spoiler wins. Otherwise, Spoiler then chooses $a_{i+1} \in R^{\mathscr{A}}(a_i)$, and $b_{i+1} := \theta_i(a_{i+1})$. Duplicator wins this round if for all P, $P^{\mathscr{A}}(a_{i+1})$ iff $P^{\mathscr{B}}(b_{i+1})$, otherwise Spoiler wins.

This game is evidently analogous to the bijection game we encountered previously.

Proposition 5.5. *The following are equivalent:*

1. *There is a winning strategy for Duplicator in the k-round graded bisimulation game between $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$.*

2. *$a \sim_k^{\mathbf{g}} b$.*

3. *$(\mathscr{A}, a) \cong_k^{\mathbb{M}} (\mathscr{B}, b)$.*

Proof. $(1) \Rightarrow (2)$ We prove the statement by induction on k. For the base case $k = 0$, by (1), Duplicator has a winning strategy for the 0-round game, so for all P, $P^{\mathscr{A}}(a)$ iff $P^{\mathscr{B}}(b)$, hence $a \sim_0^{\mathbf{g}} b$. Let $k' = k+1$ and assume that Duplicator has a winning strategy in the k'-round graded bisimulation game between $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$. By the winning condition of this strategy, for all P, $P^{\mathscr{A}}(a)$ iff $P^{\mathscr{B}}(b)$. Moreover, for all R_α, Duplicator responds to Spoiler choosing R_α in the first round of the game with a bijection $\theta_1 : R_\alpha^{\mathscr{A}}(a) \cong R_\alpha^{\mathscr{B}}(b)$. It follows that for all $a' \in R_\alpha^{\mathscr{A}}(a)$, Duplicator has a winning strategy in the k-round game from $(\mathscr{A}, a')$ to $(\mathscr{B}, \theta_1(a'))$. This fact together with the inductive hypothesis shows that $a' \sim_k^{\mathbf{g}} \theta_1(a')$. Hence, $a \sim_{k'}^{\mathbf{g}} b$.

$(2) \Rightarrow (3)$ We prove the statement by induction on k. For the base case $k = 0$, we construct the coextension morphisms $f^* : \mathbb{M}_0(\mathscr{A}, a) \to \mathbb{M}_0(\mathscr{B}, b)$ by $f^*[a] = [b]$ and $g^* : \mathbb{M}_0(\mathscr{B}, b) \to \mathbb{M}_0(\mathscr{A}, a)$ by $g^*[b] = [a]$. By (2), $a \sim_0^{\mathbf{g}} b$, so f and g preserve unary relations P. Binary relations R_α are vacuously preserved since $\mathbb{M}_0(\mathscr{A}, a)$ and $\mathbb{M}_0(\mathscr{B}, b)$ only have one element with no self-transitions. By definition of ϵ, $f \circ g^*[b] = b = \epsilon_{\mathscr{B}}[b]$ and $g \circ f^*[a] = a = \epsilon_{\mathscr{A}}[a]$. Therefore, f and g are inverses in the coKleisli category. Let $k' = k+1$ and assume that $a \sim_{k'}^{\mathbf{g}} b$. We need to define coextension morphisms $f^* : \mathbb{M}_k(\mathscr{A}, a) \to \mathbb{M}_k(\mathscr{B}, b)$ and $g^* : \mathbb{M}_k(\mathscr{B}, b) \to \mathbb{M}_k(\mathscr{A}, a)$. Since $a \sim_{k'}^{\mathbf{g}} b$, for every R_α, there exists a bijection $\theta : R_\alpha^{\mathscr{A}}(a) \cong R_\alpha^{\mathscr{B}}(b)$ such that for every $a' \in R_\alpha^{\mathscr{A}}(a)$, $a' \sim_k^{\mathbf{g}} \theta(a')$. For every $a' \in R_\alpha^{\mathscr{A}}(a)$, let $b' = \theta(a')$. We can conclude $(\mathscr{A}, a') \cong_k^{\mathbb{M}} (\mathscr{B}, b')$ from the inductive hypothesis and $a' \sim_k^{\mathbf{g}} b'$. Hence, there exists a coKleisli isomorphism given by the pair $f_{a',\alpha}^* : \mathbb{M}_k(\mathscr{A}, a') \to \mathbb{M}_k(\mathscr{B}, b')$ and $g_{b',\alpha}^* : \mathbb{M}_k(\mathscr{B}, b') \to \mathbb{M}_k(\mathscr{A}, a')$. For every R_α and $a' \in R_\alpha^{\mathscr{A}}(a)$, we define f^* as $f^*([a, \alpha, a']) = [b, \alpha, b']$ and $f^*([a, \alpha]s) = [b, \alpha]f_{a',\alpha}^*(s)$ (by $s \in \mathbb{M}_k(\mathscr{A}, a')$, a' is the first element of s). Similarly, for $\theta(a') = b'$ we define g^* as $g^*([b, \alpha, b']) = [a, \alpha, a']$ and $g^*([b, \alpha]t) = [a, \alpha]g_{b',\alpha}^*(t)$. Finally, let $f^*[a] = [b]$ and $g^*[b] = [a]$. We must show that f and g are inverses. To show that, $g^* \circ f^* = \mathsf{id}_{\mathbb{M}_{k'}(\mathscr{A}, a)}$, there are three cases:

1. Suppose $s = [a]$, then $g \circ f^*[a] = g[b] = [a]$

2. Suppose $s = [a, \alpha, a']$ for some R_α and $a' \in R_\alpha^{\mathscr{A}}(a)$:

$$g^* \circ f^*[a, \alpha, a'] = g^*[b, \alpha, b']$$
$$= [a, \alpha, a']$$

3. Suppose $s = [a, \alpha]s'$ for some R_α, $a' \in R_\alpha^{\mathscr{A}}(a)$, and $\theta(a') = b'$:

$$g^* \circ f^*([a, \alpha]s') = g^*([b, \alpha]f_{a',\alpha}^*(s'))$$
$$= [a, \alpha](g_{b',\alpha}^*(f_{a',\alpha}^*(s'))$$
$$= [a, \alpha](\mathsf{id}_{\mathbb{M}_k(\mathscr{A},a')}(s'))$$
$$= [a, \alpha]s'$$

The proof that $f^* \circ g^* = \mathsf{id}_{\mathscr{B}}$ is similar. Therefore, f is a coKleisli isomorphism, so $(\mathscr{A}, a) \cong_k^{\mathbb{M}} (\mathscr{B}, b)$.

$(3) \Rightarrow (1)$ Assume $f : \mathbb{M}_k(\mathscr{A}, a) \to (\mathscr{B}, b)$ is a coKleisli isomorphism. Let $a_0 = a$, $s_0 = [a]$ and $b_0 = b$. The fact that $P^{\mathscr{A}}(a_0) \Leftrightarrow P^{\mathscr{B}}(b_0)$ follows from f being a coKleisli isomorphism. At round $i + 1$, if Spoiler chooses R_α, Duplicator chooses the bijection $\theta_{s_i,\alpha} : R_\alpha^{\mathscr{A}}(a_i) \cong R_\alpha^{\mathscr{B}}(b_i)$ defined by $\theta_{s_i,\alpha}(a') = f(s_i[\alpha, a'])$. Spoiler then chooses an $a_{i+1} \in R_\alpha^{\mathscr{A}}(a_i)$; define $b_{i+1} = \theta_{s_i,\alpha}(a_{i+1})$ and $s_{i+1} = s_i[\alpha, a_{i+1}]$. We must show that this determines a winning move for Duplicator, i.e. for all unary $P \in \sigma$, $P^{\mathscr{A}}(a_{i+1}) \Leftrightarrow P^{\mathscr{B}}(b_{i+1})$. By f being a coKleisli isomorphism and the definition of $P^{\mathbb{M}_k(\mathscr{A},a)}$, we obtain the following chain of equivalences $P^{\mathscr{A}}(a_{i+1}) \Leftrightarrow P^{\mathbb{M}_k(\mathscr{A},a)}(s_i[\alpha, a_{i+1}]) \Leftrightarrow P^{\mathscr{B}}(f(s_i[\alpha, a_{i+1}])) \Leftrightarrow P^{\mathscr{B}}(\theta_{s_i,\alpha}(a_{i+1})) \Leftrightarrow P^{\mathscr{B}}(b_{i+1})$. $\square$

Theorem 5.6. *1. For all Kripke structures $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$:*

$$(\mathscr{A}, a) \equiv^{\exists \mathbb{M}_k} (\mathscr{B}, b) \iff (\mathscr{A}, a) \rightleftarrows_k^{\mathbb{M}} (\mathscr{B}, b).$$

2. For all image-finite Kripke structures $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$:

$$(\mathscr{A}, a) \equiv^{\mathbb{M}_k(\#)} (\mathscr{B}, b) \iff (\mathscr{A}, a) \cong_k^{\mathbb{M}} (\mathscr{B}, b).$$

Proof. (1) It is a standard result (see e.g. [10, Theorem 2.78]) that $(\mathscr{A}, a) \equiv^{\exists \mathbb{M}_k} (\mathscr{B}, b)$ iff $a \preceq_k b$ and $b \preceq_k a$. Combining this with Proposition 3.6 yields the result.

(2) It is shown in [17] and [6, Proposition 4.11] that $(\mathscr{A}, a) \equiv^{\mathbb{M}_k(\#)} (\mathscr{B}, b)$ iff $a \sim_k^{\mathsf{g}} b$. Combining this with Proposition 5.5 yields the result. $\square$

6 Coalgebras and combinatorial parameters

Another fundamental aspect of comonads is that they have an associated notion of *coalgebra*. A coalgebra for a comonad (G, ε, δ) is a morphism $\alpha : A \to GA$ such that the following diagrams commute:

$$
\begin{array}{ccc}
A & \xrightarrow{\;\alpha\;} & GA \\
{\scriptstyle \alpha}\downarrow & & \downarrow{\scriptstyle \delta_A} \\
GA & \xrightarrow{\;G\alpha\;} & G^2 A
\end{array}
\qquad\qquad
\begin{array}{ccc}
A & \xrightarrow{\;\alpha\;} & GA \\
& {\scriptstyle \mathrm{id}_A}\searrow & \downarrow{\scriptstyle \varepsilon_{\mathscr{A}}} \\
& & A
\end{array}
$$

Our use of indexed comonads $\mathbb{C}_k$ opens up a new kind of question for coalgebras. Given a structure $\mathscr{A}$, we can ask: what is the least value of k such that a $\mathbb{C}_k$-coalgebra exists on $\mathscr{A}$? We call this the *coalgebra number* of $\mathscr{A}$. We shall find that for each of our comonads, the coalgebra number is a significant combinatorial parameter of the structure.

6.1 The Ehrenfeucht-Fraïssé comonad and tree-depth

A graph is $G = (V, \frown)$, where V is the set of vertices, and $\frown$ is the adjacency relation, which is symmetric and irreflexive. A *forest cover* for G is a forest $(F, \leq)$ such that $V \subseteq F$, and if $v \frown v'$, then $v{\uparrow}v'$. The *tree-depth* $\mathsf{td}(G)$ is defined to be $\min_F \mathsf{ht}(F)$, where F ranges over forest covers of G.[3] It is clear that we can restrict to forest covers of the form $(V, \leq)$, since given a forest cover $(F, \leq)$ of $G = (V, \frown)$, $(V, \leq \cap\, V^2)$ is also a forest cover of G, and $\mathsf{ht}(V) \leq \mathsf{ht}(F)$. Henceforth, by forest covers of G we shall mean those with universe V.

Given a σ-structure $\mathscr{A}$, the [Gaifman graph] $\mathcal{G}(\mathscr{A})$ is $(A, \frown)$, where $a \frown a'$ iff for some relation $R \in \sigma$, for some $(a_1, \ldots, a_n) \in R^{\mathscr{A}}$, $a = a_i$, $a' = a_j$, $a \neq a'$. The *tree-depth of* $\mathscr{A}$ is $\mathsf{td}(\mathcal{G}(\mathscr{A}))$.

Theorem 6.1. *Let $\mathscr{A}$ be a finite σ-structure, and $k > 0$. There is a bijective correspondence between*

1. *$\mathbb{E}_k$-coalgebras $\alpha : \mathscr{A} \to \mathbb{E}_k \mathscr{A}$.*

2. *Forest covers of $\mathcal{G}(\mathscr{A})$ of height $\leq k$.*

Proof. Suppose that $\alpha : \mathscr{A} \to \mathbb{E}_k \mathscr{A}$ is a coalgebra. For $a \in A$, let $\alpha(a) = [a_1, \ldots, a_j]$, then the action of the comultiplication $\delta_{\mathscr{A}}$ on $\alpha(a)$ is $[[a_1], [a_1, a_2], \ldots, [a_1, \ldots, a_j]]$. Hence, the first coalgebra equation states that $\alpha(a_i) = [a_1, \ldots, a_i]$, $1 \leq i \leq j$. The second states that $a_j = a$. Thus $\alpha : A \to A^{\leq k}$ is an injective map whose image is a prefix-closed subset of $A^{\leq k}$. Defining $a \leq a'$ iff $\alpha(a) \sqsubseteq \alpha(a')$ yields a forest order on A, of height $\leq k$. If $a \frown a'$ in $\mathcal{G}(\mathscr{A})$, for some $a_1, \ldots, a_n$ with $a = a_i$, $a' = a_j$, we have $R^{\mathscr{A}}(a_1, \ldots, a_n)$. Since α is a homomorphism, we must have $R^{\mathbb{E}_k \mathscr{A}}(\alpha(a_1), \ldots, \alpha(a_n))$. By the pairwise comparability condition in the definition of $R^{\mathbb{E}_k \mathscr{A}}$, $\alpha(a_i){\uparrow}\alpha(a_j)$, and so $a_i{\uparrow}a_j$. Thus $(A, \leq)$ is a forest cover of $\mathscr{A}$, of height $\leq k$.

Conversely, given such a forest cover $(A, \leq)$, for each $a \in A$, its predecessors form a chain $a_1 < \cdots < a_j$, with $a_j = a$, and $j \leq k$. We define

---

[3]We formulate this notion in order-theoretic rather than graph-theoretic language, but it is equivalent to the definition in [37].

$\alpha(a) = [a_1, \ldots, a_j]$, which yields a map $\alpha : A \to A^{\leq k}$, which evidently satisfies the coalgebra equations. If $R^{\mathscr{A}}(a_1, \ldots, a_n)$, then since $(A, \leq)$ is a forest cover, we must have $a_i \uparrow a_j$ for all i, j, and hence $\alpha(a_i) \uparrow \alpha(a_j)$. Thus α is a homomorphism. $\quad\square$

We write $\kappa^{\mathbb{E}}(\mathscr{A})$ for the coalgebra number of $\mathscr{A}$ with respect to the Ehrenfeucht-Fraïssé comonad.

Theorem 6.2. *For all finite structures $\mathscr{A}$: $\mathsf{td}(\mathscr{A}) = \kappa^{\mathbb{E}}(\mathscr{A})$.*

Proof. By Theorem 6.1, for all $k > 0$, $\mathsf{td}(\mathscr{A}) \leq k$ iff $\kappa^{\mathbb{E}}(\mathscr{A}) \leq k$. $\quad\square$

6.2 The pebbling comonad and tree-width

We review the notions of tree decompositions and tree-width. A tree $(T, \leq)$ is a forest with a least element (the root). A tree is easily seen to be a meet-semilattice: every pair of elements x, x' has a greatest lower bound $x \wedge x'$ (the greatest common ancestor). The path from x to x' is the set $\mathsf{path}(x, x') := [x \wedge x', x] \cup [x \wedge x', x']$, where we use interval notation: $[y, y'] := \{z \in T \mid y \leq z \leq y'\}$.

A *tree-decomposition* of a graph $G = (V, \frown)$ is a tree $(T, \leq)$ together with a labelling function $\lambda : T \to \mathscr{P}(V)$ satisfying the following conditions:

- (TD1) for all $v \in V$, for some $x \in T$, $v \in \lambda(x)$;

- (TD2) if $v \frown v'$, then for some $x \in T$, $\{v, v'\} \subseteq \lambda(x)$;

- (TD3) if $v \in \lambda(x) \cap \lambda(x')$, then for all $y \in \mathsf{path}(x, x')$, $v \in \lambda(y)$.

The *width* of a tree decomposition is given by $\max_{x \in T} |\lambda(x)| - 1$. We define the *tree-width* $\mathsf{tw}(G)$ of a graph G as $\min_T \mathsf{width}(T)$, where T ranges over tree decompositions of G.

We shall now give an alternative formulation of tree-width which will provide a useful bridge to the coalgebraic characterization. It is also interesting in its own right: it clarifies the relationship between tree-width and tree-depth, and shows how pebbling arises naturally in connection with tree-width.

A *k-pebble forest cover* for a graph $G = (V, \frown)$ is a forest cover $(V, \leq)$ together with a pebbling function $p : V \to \mathbf{k}$ such that, if $v \frown v'$ with $v \leq v'$, then for all $w \in (v, v']$, $p(v) \neq p(w)$.

The following result is implicit in [3], but it seems worthwhile to set it out more clearly.

Theorem 6.3. *Let G be a finite graph. The following are equivalent:*

1. *G has a tree decomposition of width $< k$.*

2. *G has a k-pebble forest cover.*

Proof. (1) $\Rightarrow$ (2). Assume that $G = (V, \frown)$ has a tree decomposition $(T, \leq, \lambda)$ of width $< k$. We say that a tree decomposition is *orderly* if it has the following property: for all $x \in T$, there is at most one $v \in \lambda(x)$ such that for all $x' < x$, $v \notin \lambda(x')$. By Definition 13.1.11 of [28], *nice* tree decompositions are orderly. Moreover, by Lemma 13.1.12 of [28], the existence of a tree decomposition implies the existence of a nice tree decomposition of the same width. Hence, the existence of a tree decomposition implies the existence of an orderly one of the same width, and without loss of generality we can assume that the given tree decomposition is orderly.

For any $v \in V$, the set of $x \in T$ such that $v \in \lambda(x)$ is non-empty by (TD1), and closed under meets by (TD3). Since T is a tree, this implies that this set has a least element $\tau(v)$. This defines a function $\tau : V \to T$. The fact that the tree decomposition is orderly implies that τ is injective. We can define an order on V by $v \leq v'$ iff $\tau(v) \leq \tau(v')$. This is isomorphic to a sub-poset of T, and hence is a forest order.

We define $p : V \to \mathbf{k}$ by induction on this order. Assuming $p(v')$ is defined for all $v' < v$, we consider $\tau(v)$. Since the tree decomposition is orderly, this means in particular that $p(v')$ is defined for all $v' \in S := \lambda(\tau(v)) \setminus \{v\}$. Since the decomposition is of width $< k$, we must have $|S| < k$. We set $p(v) := \min(\mathbf{k} \setminus \{p(v') \mid v' \in S\})$.

To verify that $(V, \leq)$ is a forest cover, suppose that $v \frown v'$. By (TD2), for some $x \in T$, $\{v, v'\} \subseteq \lambda(x)$. We have $\tau(v) \leq x \geq \tau(v')$, and since T is a tree, we must have $\tau(v) \uparrow \tau(v')$, whence $v \uparrow v'$.

Finally, we must verify the condition on the pebbling function p. Suppose that $v \frown v'$, and $v < w \leq v'$. Since $v \frown v'$, for some x, $\{v, v'\} \subseteq \lambda(x)$. But then $\tau(v) < \tau(w) \leq \tau(v') \leq x$. Since $v \in \lambda(\tau(v)) \cap \lambda(x)$, by (TD3), $v \in \lambda(\tau(w))$. By construction of the pebbling function, this implies $p(v) \neq p(w)$.

(2) $\Rightarrow$ (1). Suppose that $(V, \leq, p)$ is a k-pebble forest cover of G. We define a tree $T = V_\bot$ by adjoining a least element $\bot$ to V. We say that v' is an *active predecessor* of v if $v' \leq v$, and for all $w \in (v', v]$, $p(v') \neq p(w)$. We define the labelling function by setting $\lambda(v)$ to be the set of active predecessors of v; $\lambda(\bot) := \varnothing$. Since $p|_{\lambda(v)}$ is injective, $|\lambda(v)| \leq k$.

We verify the tree decomposition conditions. (TD1) holds, since $v \in \lambda(v)$. (TD2) If $v \frown v'$, then $v \uparrow v'$. Suppose $v \leq v'$. Then v is an active predecessor of v', and $\{v, v'\} \subseteq \lambda(v')$. (TD3) Suppose $v \in \lambda(v_1) \cap \lambda(v_2)$. Then v is an active predecessor of both v_1 and v_2. This implies that for all $w \in \mathsf{path}(v_1, v_2)$, v is an active predecessor of w, and hence $v \in \lambda(w)$. $\square$

Theorem 6.4. *Let $\mathscr{A}$ be a finite σ-structure. There is a bijective correspondence between:*

1. *$\mathbb{P}_k$-coalgebras $\alpha : \mathscr{A} \to \mathbb{P}_k \mathscr{A}$*

2. *k-pebble forest covers of $\mathcal{G}(\mathscr{A})$.*

Proof. See [3, Theorem 6]. □

We write $\kappa^{\mathbb{P}}(\mathscr{A})$ for the coalgebra number of $\mathscr{A}$ with respect to the pebbling comonad.

Theorem 6.5. *For all finite structures $\mathscr{A}$: $\mathsf{tw}(\mathscr{A}) = \kappa^{\mathbb{P}}(\mathscr{A}) - 1$.*

6.3 The modal comonad and synchronization tree height

Let $\mathscr{A}$ be a Kripke structure. It will be convenient to write labelled transitions $a \xrightarrow{\alpha} a'$ for $R_\alpha(a, a')$. Given $a \in A$, the submodel generated by a, denoted $\mathscr{S}_a$, is obtained by restricting the universe to the set S_a of a' such that there is a path $a \xrightarrow{\alpha_1} \cdots \xrightarrow{\alpha_k} a'$. This submodel forms a *synchronization tree* [34] if for all a', there is a unique such path. The height of such a tree is the maximum length of any path from the root a.

Proposition 6.6. *Let $\mathscr{A}$ be a Kripke structure, with $a \in A$. The following are equivalent:*

1. *There is a coalgebra $\alpha : (\mathscr{S}_a, a) \to \mathbb{M}_k(\mathscr{S}_a, a)$.*

2. *$\mathscr{S}_a$ is a synchronization tree of height $\leq k$.*

Proof. Suppose there is a coalgebra $\alpha : (\mathscr{S}_a, a) \to \mathbb{M}_k(\mathscr{S}_a, a)$. For $a' \in S_a$, let $\alpha(a') = [a_0, \alpha_1, a_1, \ldots, \alpha_j, a_j]$ where $a_0 = a$. The first coalgebra equation states that $\alpha(a_i) = [a_0, \alpha_1 \ldots, \alpha_i, a_i]$ for $0 \leq i \leq j$. The second coalgebra equation states that $a_j = a'$. Therefore, α is injective. By injectivity, the path of transitions determined by $\alpha(a')$, i.e. $a \xrightarrow{\alpha_1} a_1 \cdots \xrightarrow{\alpha_j} a'$ is unique. Hence, every $a' \in S_a$ has a unique path from a, so the submodel generated by a is a synchronization tree. Since $j \leq k$, the height of the tree is at most k.

Conversely, suppose $\mathscr{S}_a$ is a synchronization tree of height $\leq k$, then for every a' in this submodel, there exists a unique path of transitions $a \xrightarrow{\alpha_1} \cdots \xrightarrow{\alpha_j} a'$. We can define the morphism $\alpha : (\mathscr{S}_a, a) \to \mathbb{M}_k(\mathscr{S}_a, a)$ as $\alpha(a') = [a, \alpha_1, a_1, \ldots, \alpha_j, a']$. Since the path of transitions is unique, the first coalgebra equation is satisfied. Moreover, since the last element of $\alpha(a')$ is a', the second coalgebra equation is satisfied. □

We define the modal depth $\mathsf{md}(\mathscr{A}, a) = k$ if the submodel $\mathscr{S}_a$ generated by a is a synchronization tree of height k.

Theorem 6.7. *Let $\mathscr{A}$ be a Kripke structure, and $a \in A$ be such that the submodel generated by a is a synchronization tree of finite height. Then $\mathsf{md}(\mathscr{A}, a) = \kappa^{\mathbb{M}}(\mathscr{A}, a)$.*

Note the conditional nature of this result, which contrasts with those for the other comonads. The modal comonad is defined in such a way that the universe $\mathbb{M}_k(A, a)$ reflects information about the possible transitions. Thus having a coalgebra at all, regardless of the value of the resource parameter, is a strong constraint on the structure of the transition system.

7 Characterization of Rossman-type equivalences

As a simple application of the results of the previous section, we show how they yield characterizations of some approximation preorders on structures, and their associated equivalences.

We begin with a completely general result. Given objects A, B of a category $\mathcal{C}$, we shall use the notation $A \to B$ to mean that there exists a morphism from A to B in $\mathcal{C}$.

Proposition 7.1. *Let G be a comonad on a category $\mathcal{C}$. For all objects A, B of $\mathcal{C}$, the following are equivalent:*

1. *$GA \to B$*

2. *For all G-coalgebras $\alpha : C \to GC$, $C \to A \Rightarrow C \to B$*

Proof. Suppose $f : GA \to B$, $\alpha : C \to GC$ is a coalgebra, and $g : C \to A$. Then $f \circ Gg \circ \alpha : C \to B$.

Conversely, suppose that for all coalgebras $\alpha : C \to GC$, $C \to A$ implies $C \to B$. Then since $\delta_A : GA \to GGA$ is a coalgebra, and $\varepsilon_A : GA \to A$, we conclude that $GA \to B$. □

Rossman defined the following preorder in [40]: $\mathscr{A} \to_k^{\mathbb{E}} \mathscr{B}$ iff for all structures $\mathscr{C}$ with $\mathsf{td}(\mathscr{C}) \leq k$, $\mathscr{C} \to \mathscr{A} \Rightarrow \mathscr{C} \to \mathscr{B}$. The associated equivalence $\rightleftarrows_k$ plays a major role in [40].

We can similarly define $\mathscr{A} \to_k^{\mathbb{P}} \mathscr{B}$ in terms of treewidth: $\mathscr{A} \to_k^{\mathbb{P}} \mathscr{B}$ iff for all structures $\mathscr{C}$ with $\mathsf{tw}(\mathscr{C}) < k$, $\mathscr{C} \to \mathscr{A} \Rightarrow \mathscr{C} \to \mathscr{B}$ [4].

Finally, we can define $(\mathscr{A}, a) \to_k^{\mathbb{M}} (\mathscr{B}, b)$ iff for all synchronization trees $(\mathscr{C}, c)$ with $\mathsf{md}(\mathscr{C}, c) \leq k$, $(\mathscr{C}, c) \to (\mathscr{A}, a) \Rightarrow (\mathscr{C}, c) \to (\mathscr{B}, b)$.

As an immediate consequence of Proposition 7.1 and the results of the previous section, we obtain:

Proposition 7.2. *For all structures A, B in $\mathcal{R}(\sigma)$:*

1. *$\mathscr{A} \to_k^{\mathbb{E}} \mathscr{B} \iff \mathbb{E}_k\mathscr{A} \to \mathscr{B}$*

2. *$\mathscr{A} \to_k^{\mathbb{P}} \mathscr{B} \iff \mathbb{P}_k\mathscr{A} \to \mathscr{B}$*

3. *$(\mathscr{A}, a) \to_k^{\mathbb{M}} (\mathscr{B}, b) \iff \mathbb{M}_k(\mathscr{A}, a) \to (\mathscr{B}, b)$ (when $\mathscr{A}, \mathscr{B}$ are Kripke structures)*

---

[4] We use strict inequality, in this case, due to the fact that tree-width is one less than the $\mathbb{P}_k$-coalgebra number.

8 Coalgebras and conjunctive queries

We shall now connect coalgebras with logic. We will show that they are closely related to *conjunctive queries*. We shall focus on boolean conjunctive queries, *i.e.* closed formulas built from atomic formulae using only conjunction and existential quantification. We allow the empty conjunction $\top$. Given a finite structure $\mathscr{A}$ with universe $A = \{a_1, \ldots, a_n\}$, the canonical conjunctive query for $\mathscr{A}$ is a formula $Q^{\mathscr{A}} := \exists v_1, \ldots, v_n. \bigwedge \{R(\vec{v}) \mid R \in \sigma, \vec{a} \in R^{\mathscr{A}}\}$. Here the correspondence $v_i \leftrightarrow a_i$ uses the linear ordering on the universe implicitly given by the enumeration.

The key property of the canonical conjunctive query is the following [12]:

Theorem 8.1. *For all finite structures $\mathscr{A}$, $\mathscr{B}$, $\mathscr{A} \to \mathscr{B}$ iff $\mathscr{B} \models Q^{\mathscr{A}}$.*

We now consider a number of rewrite rules on formulas:

$(R1)$ Associative-Commutative-Identity rewriting of conjunctions
$(R2)$ $\exists v. (\varphi \wedge \psi) \rightsquigarrow (\exists v. \varphi) \wedge \psi$ $(v \notin \mathsf{FV}(\psi))$
$(R3)$ $\exists v. \exists w. \varphi \rightsquigarrow \exists w. \exists v. \varphi$
$(R4)$ $\exists v. \varphi \rightsquigarrow \exists w. \phi[w/v]$

Note that $(R1)$ amounts to taking formulas as a commutative monoid under the binary operation of conjunction, with $\top$ as the identity.

As usual, we can use a rule $\theta \rightsquigarrow \chi$ to rewrite a formula $C[\theta]$ to $C[\chi]$. We write $\varphi \twoheadrightarrow^{\mathbb{E}} \psi$ if φ can be rewritten to ψ in some finite number of steps using rules (R1)–(R3).

Theorem 8.2. *Let $\mathscr{A}$ be a finite structure, and $Q^{\mathscr{A}}$ its canonical conjunctive query. The following are equivalent:*

1. $\mathscr{A}$ has a coalgebra $\mathscr{A} \to \mathbb{E}_k \mathscr{A}$

2. $Q^{\mathscr{A}} \twoheadrightarrow^{\mathbb{E}} \varphi$, where φ has quantifier rank $\leq k$.

Proof. Assuming (2), we can define a forest order on A, with $a_i < a_j$ if v_j occurs in the scope of v_i in φ. Since the quantifier rank of φ is $\leq k$, so is the height of this order. Moreover, condition (E) is obviously satisfied by the order induced by $Q^{\mathscr{A}}$, and is preserved by applications of the rules (R1)–(R3). In particular, in applying (R2), the free variable condition implies that no elements which become incomparable in the corresponding order are adjacent in the Gaifman graph. By induction on the length of the rewrite sequence, we conclude that the order induced by φ is a forest cover, and hence, by Theorem 6.1, determines an $\mathbb{E}_k$-coalgebra on $\mathscr{A}$.

For the converse, we argue by induction on the cardinality of A, or equivalently on the number n of variables in $Q^{\mathscr{A}}$. We prove the following statement $\Phi(n)$ by induction on n:

For all finite relational vocabularies σ, and for all σ-structures $\mathscr{A}$ of cardinality n, if $\mathscr{A}$ has a forest cover of height k, then $Q^{\mathscr{A}} \twoheadrightarrow^{\mathbb{E}} \varphi$, where the quantifier rank of φ is $\leq k$.

The base case $n = 1$ is trivial. Now consider

$$Q^{\mathscr{A}} = \exists v_1, \ldots, v_j. \exists v_{j+1}, \ldots, v_n. \bigwedge_{i \in I} R_i(\vec{v}_i).$$

We assume that the variables are in an order which linearizes the given forest order on A. In particular, $v_{j+1}, \ldots, v_n$ are the leaves (maximal elements) of maximum height k in the forest order. Note that if v and v' are leaves, they cannot be adjacent in the Gaifman graph of $\mathscr{A}$, since this would violate the forest cover condition. Hence we can partition I as $I = I_{j+1} \sqcup \cdots I_n \sqcup L$, where I_p is the set of indices labelling atomic formulas in which v_p occurs, $p = j + 1, \ldots, n$. Using (R1) and (R2), we can rewrite $Q^{\mathscr{A}}$ to

$$\varphi := \exists v_1, \ldots, v_j. \left(\bigwedge_{p=j+1}^{n} \chi_p \right) \wedge \theta$$

where $\chi_p := \exists v_p. \bigwedge_{j \in I_p} R_j(\vec{v}_j)$ and $\theta := \bigwedge_{l \in L} R_l(\vec{v}_l)$. We now define

$$\varphi' := \exists v_1, \ldots, v_j. \left(\bigwedge_{p=j+1}^{n} R_p(\vec{w}) \right) \wedge \theta,$$

where R_p is a new relation symbol, and $\vec{w}_p$ lists all the variables other than v_p occurring in χ_p. This formula φ' is the canonical conjunctive query for a new structure $\mathscr{A}'$, of cardinality $j < n$. There are many new adjacencies in the Gaifman graph of $\mathscr{A}'$ compared with that for $\mathscr{A}$. However, note that, since (the element of A labelled by) each v_p was maximal in the given forest cover of $\mathscr{A}$, the set of variables adjacent to v_p must be linearly ordered in that forest cover. Hence the restriction of the forest cover on $\mathscr{A}$ to $\{v_1, \ldots, v_j\}$ is a forest cover on $\mathscr{A}'$, of height $k-1$. Our induction hypothesis can be applied to this forest cover on $\mathscr{A}'$ as a σ'-structure, where $\sigma' = \sigma \cup \{R_p \mid p = j+1, \ldots, n\}$, yielding a rewrite $\varphi' \twoheadrightarrow^{\mathbb{E}} \psi'$, where ψ' has quantifier rank $m < k$. We can perform the same rewrite steps to obtain $\varphi \twoheadrightarrow^{\mathbb{E}} \psi$, where ψ results from ψ' by replacing each $R_p(\vec{w})$ by χ_p. We have $Q^{\mathscr{A}} \twoheadrightarrow^{\mathbb{E}} \varphi \twoheadrightarrow^{\mathbb{E}} \psi$. Moreover, the quantifier rank of ψ is $m + 1 \leq k$, as required. $\qquad\square$

We can think of (R1)–(R3) as encoding a non-deterministic algorithm for computing the minimum quantifier rank for $Q^{\mathscr{A}}$ (and hence the tree-depth of $\mathscr{A}$). We use (R3) to guess an order on the variables, and then apply (R1) and (R2) to the quantifier prefix from the inside out. The above argument shows that this algorithm does, for some choice of order, find the minimum quantifier rank.

We can give a very similar analysis for the pebbling case. We define $\varphi \twoheadrightarrow^{\mathbb{P}} \psi$ if φ can be rewritten to ψ in some finite number of steps using rules (R1), (R2), and (R4). This set of rules is widely used in query optimization [16, 43].

Theorem 8.3. *Let $\mathscr{A}$ be a finite structure, and $Q^{\mathscr{A}}$ its canonical conjunctive query. The following are equivalent:*

1. *$\mathscr{A}$ has a coalgebra $\mathscr{A} \to \mathbb{P}_k \mathscr{A}$*

2. *$Q^{\mathscr{A}} \twoheadrightarrow^{\mathbb{P}} \varphi$, where φ has number of variables $\leq k$.*

Proof. This is the statement of Theorem 27 in [3]. $\square$

For the modal case, we moved to the category of pointed Kripke structures. Syntactically, this corresponds to modifying the canonical conjunctive query to have one free variable; we write this as $Q^a(x)$. Note that we can rearrange the block of existential quantifiers in $Q^{\mathscr{A}}$ using rule (R3) so that $\exists x$ is the leftmost existential quantifier. Hence, we could write $Q^{\mathscr{A}}$ as $\exists x Q^a(x)$, so $(\mathscr{A}, a) \vDash Q^a(x) \iff \mathscr{A} \vDash Q^{\mathscr{A}}$. We write $\varphi(x) \twoheadrightarrow^{\mathbb{M}} \psi(x)$ if φ can be rewritten to $\psi(x)$ in some finite number of steps using rules (R1)-(R4).

Theorem 8.4. *Let $\mathscr{A}$ be a finite Kripke structure with distinguished world $a \in \mathscr{A}$. The following are equivalent:*

1. *$\mathscr{S}_a$ has a coalgebra $(\mathscr{S}_a, a) \to \mathbb{M}_k(\mathscr{S}_a, a)$*

2. *$Q^a(x) \twoheadrightarrow^{\mathbb{M}} [\![\varphi]\!]$ where φ is a multi-modal logic sentence of modal-depth $\leq k$ and $[\![\cdot]\!]$ is the standard translation map into first-order logic.*

Proof. The proof is similar to the proof of 8.2, but takes into account the labels of the accessibility relations that guard quantifiers in translated modal formulas.

Assuming (2), let $\phi(x) = [\![\varphi]\!]$ be the image of modal sentence φ under the standard translation. Through repeated applications of (R4), we can assume that every variable in $\phi(x)$ is distinct. For $a' \in \mathscr{A}$, we can define $\alpha_i(a')$ by induction on the number of variables i in scope of v' where v' is the variable assigned to a' in the assignment witnessing $(\mathscr{A}, a) \vDash Q^a(x)$. For the base case of 0, we set $\alpha_0(a') = [a']$. For the inductive step, suppose v_{i+1} occurs in the scope of v', then there must exist some subformula of $\phi(x)$ of the form $\exists v_{i+1}(R_{\beta_{i+1}}(v_{i+1}, v') \wedge \chi)$. By the inductive, hypothesis, we have defined $\alpha_i(a') = s$. Hence, we define $\alpha_{i+1}(a') = [a_{i+1}, \beta_{i+1}]s$ where a_{i+1} corresponds to v_{i+1} in the assignment witnessing $(\mathscr{A}, a) \vDash Q^a(x)$. Since φ has modal depth at most k, the sequence of transitions in $\alpha_k(a')$ has length at most k (where we count one for each labelled transition). Finally, we define $\alpha(a') = \alpha_i(a')$ where i is the number of variables in the scope of v'. The counit axiom states that the last element of $\alpha(a')$ is a'. By construction,

$\alpha_i(a')$ has $\alpha_j(a')$ as a suffix, with $0 \leq j < i$. By the base case, $\alpha_0(a') = [a']$, therefore the last element of $\alpha_i(a')$ is a' as desired. The comultiplication axiom states that if $\alpha(a') = [a, \beta_0, \ldots, \beta_i, a']$, then $\alpha(a_j) = [a, \beta_0, \ldots, \beta_j, a_j]$ for all $j \in \mathbf{i}$. By the construction of $\alpha_i(a')$, it follows that the variable v_j corresponding to a_j is in the scope of v'. Therefore, $\alpha_j(a_j) = \alpha(a_j) = [a, \beta_0, \ldots, \beta_j, a_j]$.

For the converse, by proposition 6.6, it suffices to show that: for every $a \in \mathscr{A}$, if $\mathscr{S}_a$ is a synchronization tree of height $\leq k$, then $Q^a(x) \twoheadrightarrow^{\mathbb{M}} \phi(x)$ where $\phi(x) = [\![\varphi]\!]$ for some modal sentence φ of depth $\leq k$. We prove this statement by induction on k. For the base case $k = 0$, we can conclude that $\mathscr{S}_a$ is such that $R_\alpha^{\mathscr{S}_a}$ is empty for every $R_\alpha \in \sigma$. Therefore, $Q^a(x)$ is a conjunction of unary atoms $P \in \sigma$ that a satisfies. We can take $\phi(x) = Q^a(x)$ and φ would be a conjunction of the propositional variables corresponding to the satisfied unary relations. For the inductive step, suppose $\mathscr{S}_a$ is a synchronization tree of height $\leq k + 1$, then we can successively apply (R1) and (R2), so that $\phi(x)$ is a conjunction of formulas of two forms:

1. $\exists y(R_\beta(x, y) \wedge Q^b(y))$, for some y corresponding to b such that $a \xrightarrow{\beta} b$

2. $P(x)$ for some unary $P \in \sigma$ such that $P^{\mathscr{A}}(a)$

Applying the inductive hypothesis to $\mathscr{S}_b$ and replacing $Q^b(y)$ with its rewrite $\chi(y)$ yields the desired $\phi(x)$. By the inductive hypothesis $\chi(y) = [\![\rho]\!]$ for some modal sentence ρ of depth $\leq k$, then φ, such that $\varphi = [\![\phi]\!]$, can be seen as a conjunction of sentences of the form $\Diamond_\beta \rho$ and propositional variables p (corresponding to each type of conjunct in $\phi(x)$). $\square$

9 Game comonads as cofree constructions

We now show that the game comonads arise from certain adjunctions in a rather natural way, building on the descriptions of the coalgebras for these comonads given in Section 6.

9.1 The Ehrenfeucht-Fraïssé adjunction

We define a *tree-ordered σ-structure* $(\mathscr{A}, \leq)$ to be a σ-structure $\mathscr{A}$ with a forest order $\leq$ on A, satisfying the following condition:

(E) if $a \frown b$ in $\mathcal{G}_\sigma(\mathscr{A})$, the Gaifman graph of $\mathscr{A}$ as a σ-structure, then $a \uparrow b$.

Note that such a structure is the same thing as a forest cover of $\mathscr{A}$ with universe A. We write the covering relation of $\leq$ as $\prec$. A morphism of tree-ordered σ-structures $f : (\mathscr{A}, \leq) \to (\mathscr{B}, \leq')$ is a σ-homomorphism $f : \mathscr{A} \to \mathscr{B}$ which maps roots to roots, and preserves the covering relation. This determines a category $\mathcal{R}^E(\sigma)$, and an evident forgetful functor $U : \mathcal{R}^E(\sigma) \to$

$\mathcal{R}(\sigma)$. For each $k > 0$, if we restrict to forest orders of height $\leq k$, we get a sub-category $\mathcal{R}_k^E(\sigma)$, and a functor $U_k : \mathcal{R}_k^E(\sigma) \to \mathcal{R}(\sigma)$.

Theorem 9.1. *For each $k > 0$, the functor U_k has a right adjoint G_k. Moreover, $\mathbb{E}_k$ is the comonad arising from this adjunction.*

Proof. We define $G_k \mathscr{A} := (\mathbb{E}_k \mathscr{A}, \sqsubseteq)$, where $\sqsubseteq$ is the prefix ordering. Note that $U_k G_k \mathscr{A} = \mathbb{E}_k \mathscr{A}$. The required universal property is the following:

$$
\begin{array}{ccc}
G_k \mathscr{A} & \qquad U_k G_k \mathscr{A} \xrightarrow{\ \varepsilon_\mathscr{A}\ } \mathscr{A} \\
\hat{f} \uparrow & \qquad U_k \hat{f} \uparrow \quad \nearrow f \\
\mathcal{T} & \qquad U_k \mathcal{T}
\end{array}
$$

Here $\mathcal{T}$ is a tree-ordered σ-structure. Given $x \in T$, we write $x_1 \prec \cdots \prec x_j = x$ for the finite chain of predecessors on x in the forest order. We define $\hat{f}(x) = [a_1, \ldots, a_j]$, where $a_i = f(x_i)$, $i = 1, \ldots, j$. To show that $\hat{f}$ is a homomorphism, suppose that $R^\mathcal{T}(x_1, \ldots, x_n)$. Then Condition (E) implies that $x_1, \ldots, x_n$ are pairwise comparable in the tree order, and hence lie on the same branch; while the fact that f is a homomorphism implies that $R^\mathscr{A}(f(x_1), \ldots, f(x_n))$. From the definition of $R^{\mathbb{E}_k \mathscr{A}}$, this implies that $R^{\mathbb{E}_k \mathscr{A}}(\hat{f}(x_1), \ldots, \hat{f}(x_n))$. Finally, unicity follows from the requirement that $\hat{f}$ has to preserve the covering relation.

Note that, given $f : \mathbb{E}_k \mathscr{A} \to \mathscr{B}$, this is $f : U_k G_k \mathscr{A} \to \mathscr{B}$, and we can define the coKleisli extension by $f^* := U_k \hat{f} : \mathbb{E}_k \mathscr{A} \to \mathbb{E}_k \mathscr{B}$. $\qquad \square$

Theorem 9.2. *The coalgebras $\alpha : \mathscr{A} \to \mathbb{E}_k \mathscr{A}$ are exactly the morphisms of the form $U_k \widehat{\mathsf{id}_\mathscr{A}} : \mathscr{A} \to U_k G_k \mathscr{A}$ for $\mathcal{T} \in \mathcal{R}_k^E(\sigma)$ with $U_k \mathcal{T} = \mathscr{A}$.*

Proof. This is a restatement of Theorem 6.1 in the language of the adjunction $U_k \dashv G_k$. $\qquad \square$

9.2 The pebbling adjunction

We define a *k-pebble tree-ordered σ-structure* $(\mathscr{A}, \leq, p)$ to be a σ-structure $\mathscr{A}$ with a forest order $\leq$ on A, and a pebbling function $p : A \to \mathbf{k}$. In addition to condition (E), it must also satisfy the following condition:

(P) if $a \frown b$ in $\mathcal{G}_\sigma(\mathscr{A})$, and $a < b$ in the forest order, then for all $x \in (a, b]$,
$$p(a) \neq p(x).$$

Morphisms of these structures are morphisms of tree-ordered structures which additionally preserve the pebbling function. These define a category $\mathcal{R}_k^P(\sigma)$ (where k bounds the number of pebbles, rather than the height of the forest order), and there is an evident forgetful functor $V_k : \mathcal{R}_k^P(\sigma) \to \mathcal{R}(\sigma)$.

Theorem 9.3. *For each $k > 0$, the functor V_k has a right adjoint H_k. Moreover, $\mathbb{P}_k$ is the comonad arising from this adjunction.*

Proof. The proof proceeds on similar lines to that of Theorem 9.1. We define $H_k\mathscr{A} := (\mathbb{P}_k\mathscr{A}, \sqsubseteq, p)$, where $p([(p_1, a_1), \ldots, (p_j, a_j)]) = p_j$. Given $(\mathscr{B}, \leq, p)$ in $\mathcal{R}_k^P(\sigma)$, $g : \mathscr{B} \to \mathscr{A}$, and $x_1 \prec \cdots \prec x_j = x$ in $\mathscr{B}$, we define $\hat{g}(x) = [(p_1, a_1), \ldots, (p_j, a_j)]$, where $a_i = g(x_i)$ and $p_i = p(x_i)$. We now argue as in the Ehrenfeucht-Fraïssé case, using the additional condition (P) to show that $\hat{g}$ is a homomorphism, and the additional requirement of preserving the pebbling function to show unicity. $\square$

Theorem 9.4. *The coalgebras $\alpha : \mathscr{A} \to \mathbb{P}_k\mathscr{A}$ are exactly the morphisms of the form $V_k\widehat{\mathrm{id}_{\mathscr{A}}} : \mathscr{A} \to V_kH_k\mathscr{A}$ for $\mathcal{T} \in \mathcal{R}_k^P(\sigma)$ with $V_k\mathcal{T} = \mathscr{A}$.*

Proof. This is a restatement of Theorem 6.4 in the language of the adjunction $V_k \dashv H_k$. $\square$

9.3 The modal adjunction

Let σ be a modal vocabulary. We define a modal σ-structure $(\mathscr{A}, a, \leq)$ to be a σ-structure $\mathscr{A}$ with a partial order $\leq$ on A, such that $\uparrow(a) := \{a' \in A \mid a \leq a'\}$ is a tree order. This must satisfy the following condition:

$$(\mathrm{M}) \text{ for } x, y \in \uparrow(a), \ R_\alpha(x, y) \ \Rightarrow \ x \prec y.$$

A morphism of such structures is a σ-homomorphism preserving the root and the covering relation. For each $k > 0$, there is a category $\mathcal{R}_k^M(\sigma)$, with k bounding the height of the tree order, and a forgetful functor $W_k : \mathcal{R}_k^M(\sigma) \to \mathcal{R}_\star(\sigma)$, which sends $(\mathscr{A}, a, \leq)$ to $(\mathscr{A}, a)$.

Theorem 9.5. *For each $k > 0$, the functor W_k has a right adjoint, which sends $(\mathscr{A}, a)$ to $(\mathbb{M}_k\mathscr{A}, [a], \sqsubseteq)$. Moreover, $\mathbb{M}_k$ is the comonad arising from this adjunction.*

We have stated the result in this fashion for uniformity with the other cases, but it is worth noting that a simpler analysis applies in this case.[5] The modal comonad is idempotent, *i.e.* the comultiplication is an isomorphism. This means that it encodes a coreflective subcategory. The category of trees $\mathcal{R}_k^M(\sigma)$ is coreflective in the category of all pointed Kripke structures (see e.g. [47]), and the modal comonad arises from this coreflection.

9.4 *I*-morphisms and relative adjunctions

To fit *I*-morphisms, as discussed in section 4, into this picture, we shall use *relative adjunctions* [44]. Recall that given functors $J : \mathcal{B} \to \mathcal{D}$ and $L : \mathcal{C} \to \mathcal{D}$, we say that L has a J-right adjoint $R : \mathcal{B} \to \mathcal{C}$ if there is a natural isomorphism

$$\hom_\mathcal{D}(L(-), J(-)) \ \cong \ \hom_\mathcal{C}(-, R(-)).$$

---

[5]The authors are indebted to Dan Marsden for pointing this out.

In particular, if L has a *bona fide* right adjoint $R : \mathcal{D} \to \mathcal{C}$, we obtain a J-right adjoint by composing R with J.
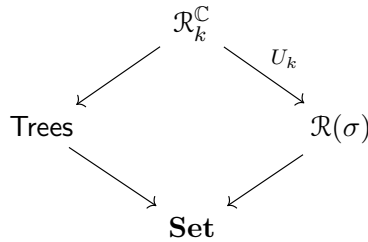
In our case, we have the functor $J : \mathcal{R}(\sigma) \to \mathcal{R}(\sigma^+)$, and we obtain relative versions of the Ehrenfeucht-Fraïssé and pebbling adjunctions, again using the fact that our constructions are uniform in the relational vocabulary, and hence can be applied to $\mathcal{R}(\sigma^+)$. These relative right adjoints give rise to the relative comonads $\mathbb{E}_k^+$ and $\mathbb{P}_k^+$.

10 Logical Equivalences II: Open pathwise embeddings and back-and-forth equivalences

For each comonad $\mathbb{C}_k$, we shall now define a "back-and-forth" equivalence $\leftrightarrow_k^{\mathbb{C}}$, intermediate between $\rightleftarrows_k^{\mathbb{C}}$ and $\cong_k^{\mathbb{C}}$, and use it to characterize the logical equivalences induced by $\mathcal{L}_k$. These back-and-forth equivalences will be more specific to "game comonads" defined on $\mathcal{R}(\sigma)$, but they will still be defined and shown to have the appropriate properties in considerable generality. We shall use a variant of the well-known notion of *open map bisimulation* of Joyal, Nielsen and Winskel [27].

Although it is folklore that Ehrenfeucht-Fraïssé equivalence is "essentially" a form of bisimulation, to our knowledge this is the first time that graded elementary equivalences have been captured in a precise common format with open map bisimulations. The key novel ingredient in our approach is the use of embeddings, and the notion of *pathwise embedding*. This allows us to capture a general notion of *property-preserving bisimulation*, which specializes to capture all the notions of interest for model comparison games.

Referring to our study of adjunctions characterizing the game comonads in Section 9, we can extract the following common structure. For each comonad $\mathbb{C}_k$ (variously $\mathbb{E}_k$, $\mathbb{P}_k$, or $\mathbb{M}_k$), we have a category $\mathcal{R}_k^{\mathbb{C}}$, and a commuting diagram of categories and functors:

$$
\begin{array}{ccc}
 & \mathcal{R}_k^{\mathbb{C}} & \\
 \swarrow & & \searrow^{U_k} \\
 \mathsf{Trees} & & \mathcal{R}(\sigma) \\
 \searrow & & \swarrow \\
 & \mathbf{Set} &
\end{array}
$$

Here Trees is the category of tree orders and maps which preserve the root and the covering relation. The unlabelled arrows are the evident faithful forgetful functors. The faithful functor U_k has a right adjoint F_k, and $\mathbb{C}_k$ arises from this adjunction.

Note that the Ehrenfeucht-Fraïssé and pebbling comonads produce *forests* of non-empty sequences rather than trees. However, it is always possible,

33

and often convenient, e.g. for the base case of an inductive proof, to turn the forests into trees by adding the empty sequence. This is justified by the fact that Forests, the category of forest orders and maps which take roots to roots, and preserve the covering relation, is equivalent to Trees.

The above description applies directly to the Ehrenfeucht-Fraïssé and pebbling cases. For the modal case, we should use $\mathcal{R}_\star(\sigma)$ rather than $\mathcal{R}(\sigma)$. In this case, the construction produces a tree directly.

We recall the notion of I-morphism from Section 4. We shall build this into the current picture by working with the J-relative versions of the Ehrenfeucht-Fraïssé and pebbling adjunctions, as in section 9.4. This means that objects of $\mathcal{R}_k^{\mathbb{C}}$ carry an interpretation of the I-relation, which morphisms of $\mathcal{R}_k^{\mathbb{C}}$ must preserve. The right adjoint functor F_k produces a relation $I^{F_k\mathscr{A}}$ by the functorial action of F_k on the identity relation $I^{\mathscr{A}}$.

A morphism e in $\mathcal{R}_k^{\mathbb{C}}$ is an *embedding* if $U_k(e)$ is an extremal mono in $\mathcal{R}(\sigma)$, *i.e.* an embedding of relational structures. We write $e : \mathscr{A} \rightarrowtail \mathscr{B}$ to indicate that e is an embedding.

We define a subcategory Paths of $\mathcal{R}_k^{\mathbb{C}}$ whose objects have tree orders which are linear, so they comprise a single branch. If P is a path, I^P is the identity relation. Morphisms of paths are embeddings. More generally, we say that $e : P \rightarrowtail \mathscr{A}$ is a *path embedding* if P is a path. A morphism $f : \mathscr{A} \to \mathscr{B}$ in $\mathcal{R}_k^{\mathbb{C}}$ is a *pathwise embedding* if for any path embedding $e : P \rightarrowtail \mathscr{A}$, $f \circ e$ is a path embedding.

We can now define what it means for a morphism $f : \mathscr{A} \to \mathscr{B}$ in $\mathcal{R}_k^{\mathbb{C}}$ to be *open*. This holds if, whenever we have a diagram

$$
\begin{array}{ccc}
P & \rightarrowtail & Q \\
\downarrow & & \downarrow \\
\mathscr{A} & \xrightarrow{f} & \mathscr{B}
\end{array}
$$

where P and Q are paths, there is an embedding $Q \rightarrowtail \mathscr{A}$ such that

$$
\begin{array}{ccc}
P & \rightarrowtail & Q \\
\downarrow & \swarrow & \downarrow \\
\mathscr{A} & \xrightarrow{f} & \mathscr{B}
\end{array}
$$

This is often referred to as the *path-lifting property*. If we think of f as witnessing a simulation of $\mathscr{A}$ by $\mathscr{B}$, path-lifting means that if we extend a given behaviour in $\mathscr{B}$ (expressed by extending the path P to Q), then we can find a matching behaviour in $\mathscr{A}$ to "cover" this extension. Thus it expresses an abstract form of the notion of "p-morphism" from modal logic [10], or of functional bisimulation.

We can now define the *back-and-forth equivalence* $\mathscr{A} \leftrightarrow_k^{\mathbb{C}} \mathscr{B}$ between structures in $\mathcal{R}(\sigma)$. This holds if there is a structure $\mathscr{R}$ in $\mathcal{R}_k^{\mathbb{C}}$, and a span

of open pathwise embeddings

$$
\begin{array}{ccc}
& \mathscr{R} & \\
& \swarrow \qquad \searrow & \\
F_k\mathscr{A} & & F_k\mathscr{B}
\end{array}
$$
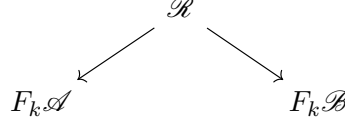
This gives us a general, structural description of back-and-forth equivalence. At the same level of generality, we shall now define a *back-and-forth game* $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ played between the structures $\mathscr{A}$ and $\mathscr{B}$, corresponding to the comonad $\mathbb{C}_k$. Positions of the game are pairs $(s,t) \in \mathbb{C}_k A \times \mathbb{C}_k B$.

We define a relation $\mathsf{W}(\mathscr{A}, \mathscr{B}) \subseteq \mathbb{C}_k A \times \mathbb{C}_k B$ as follows. A pair (s,t) is in $\mathsf{W}(\mathscr{A}, \mathscr{B})$ iff for some path P, path embeddings $e_1 : P \rightarrowtail \mathscr{A}$ and $e_2 : P \rightarrowtail \mathscr{B}$, and $p \in P$, $s = e_1(p)$ and $t = e_2(p)$. The intention is that $\mathsf{W}(\mathscr{A}, \mathscr{B})$ picks out the winning positions for Duplicator.

At the start of each round of the game, the position is specified by $(s,t) \in \mathbb{C}_k A \times \mathbb{C}_k B$. The initial position is $(\bot, \bot)$. The round proceeds as follows. Either Spoiler chooses some $s' \succ s$, and Duplicator must respond with $t' \succ t$, resulting in a new position (s', t'); or Spoiler chooses some $t'' \succ t$ and Duplicator must respond with $s'' \succ s$, resulting in (s'', t''). Duplicator wins the round if they are able to respond, and the new position is in $\mathsf{W}(\mathscr{A}, \mathscr{B})$.

Theorem 10.1. *The following are equivalent:*

1. *$\mathscr{A} \leftrightarrow_k^{\mathbb{C}} \mathscr{B}$.*

2. *There is a winning strategy for Duplicator in the $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ game.*

Proof. $(1) \Rightarrow (2)$. Firstly, consider an open pathwise embedding $q : \mathscr{R} \to F_k\mathscr{B}$. Since q is a morphism in $\mathcal{R}_k^{\mathbb{C}}$, any covering chain $\bot \prec r_1 \prec \cdots \prec r_i$ in $\mathscr{R}$ is mapped to a covering chain $\bot \prec s_1 \prec \cdots \prec s_i$ in $F_k\mathscr{B}$. Since $\bot \prec r_1 \prec \cdots \prec r_i$ is the image of a path embedding $P \rightarrowtail \mathscr{R}$, and q is a pathwise embedding, for each j with $1 \le j \le i$, $(r_j, s_j) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$.

Claim For any $s_{i+1} \succ s_i \in F_k B$, there is $r_{i+1} \succ r_i \in R$ such that $q(r_{i+1}) = s_{i+1}$.

To prove this, note that there are path embeddings $e_1 : P \rightarrowtail \mathscr{R}$ and $e_2 : Q \rightarrowtail F_k\mathscr{B}$ with images $\bot \prec r_1 \prec \cdots \prec r_i$ and $\bot \prec s_1 \prec \cdots \prec s_i \prec s_{i+1}$ respectively. Writing $Q_{\le i}$ for the truncation of Q to the preimage of $\bot \prec s_1 \prec \cdots \prec s_i$, $\mathscr{R}_i$ for the image of e_1, and $(F_k\mathscr{B})_i$ for the image of $e_2|_{Q_{\le i}}$, we have an embedding $e : P \cong \mathscr{R}_i \cong (F_k\mathscr{B})_i \cong Q_{\le i} \rightarrowtail Q$, such that $q \circ e_1 = e_2 \circ e$. Since q is open, there is an embedding $Q \rightarrowtail \mathscr{R}$ such that

$$
\begin{array}{ccc}
P & \overset{e}{\rightarrowtail} & Q \\
{\scriptstyle e_1}\downarrow & \diagup & \downarrow{\scriptstyle e_2} \\
\mathscr{R} & \underset{q}{\longrightarrow} & F_k\mathscr{B}
\end{array}
$$

We define r_{i+1} to be the image of the greatest element of Q under this embedding. Since Q is a chain, $r_{i+1} \succ r_i$. By the commutativity of the above diagram, $q(r_{i+1}) = s_{i+1}$.

Now suppose we have a span of open pathwise embeddings

$$
\begin{array}{ccc}
 & \mathscr{R} & \\
{}^{p}\swarrow & & \searrow{}^{q} \\
F_k\mathscr{A} & & F_k\mathscr{B}
\end{array}
$$

For any path embedding $P \rightarrowtail \mathscr{R}$ with image $\bot \prec r_1 \prec \cdots \prec r_i$, we have a pair $(s,t) = (p(r_i), q(r_i))$ in $\mathbb{C}_k\mathscr{A} \times \mathbb{C}_k\mathscr{B}$. Since p and q are pathwise embeddings, $(s,t) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$.

Moreover, for any $t' \succ t$, we can find $r_{i+1} \succ r_i$ and $s' \succ s$ such that $(p(r_{i+1}), q(r_{i+1})) = (s', t')$ by the above Claim. A symmetric argument applies to yield an extension for any $s' \succ s$. Thus the set of such pairs (s,t) yields a winning strategy for Duplicator in the $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ game.

(2) $\Rightarrow$ (1). Suppose we have a winning strategy for Duplicator. This determines a set $R \subseteq \mathbb{C}_k A \times \mathbb{C}_k B$ of the plays following this strategy.

R is a down-closed subset of a tree, and hence forms a tree in the induced order. It determines a substructure of the product relational structure $F_k\mathscr{A} \times F_k\mathscr{B}$, and hence an object $\mathscr{R}$ in $\mathcal{R}_k^{\mathbb{C}}$. Restricting the projections to R determines morphisms $p : \mathscr{R} \rightarrow F_k\mathscr{A}$, $q : \mathscr{R} \rightarrow F_k\mathscr{B}$.

Firstly, we show that p and q are pathwise embeddings. Given a path embedding $e : P \rightarrowtail \mathscr{R}$, $p \circ e$ is an injective homomorphism, since it is a morphism in $\mathcal{R}_k^{\mathbb{C}}$, and preserves the covering relation. It remains to show that $p \circ e$, or equivalently the restriction of p to the image of e, is strong. Given $(s_1, t_1), \ldots, (s_n, t_n)$ in the image of e, let s be the maximum of the s_i, which exists since the image of a path forms a chain. If $s = s_j$, then t_j is the maximum of the t_i. Now since $(s,t) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$, there is a path Q, and embeddings $e_1 : Q \rightarrowtail F_k\mathscr{A}$, $e_2 : Q \rightarrowtail F_k\mathscr{B}$, with $s = e_1(r)$ and $t = e_2(r)$ for some $r \in Q$. For each i, there is a unique element $r_i \in Q$ such that $e_1(r_i) = s_i$ and $e_2(r_i) = t_i$. For an n-ary relation R in σ, since e_1 and e_2 are embeddings, we have

$$
R^{F_k\mathscr{A}}(s_1, \ldots, s_n) \iff R^{Q}(r_1, \ldots, r_n) \iff R^{F_k\mathscr{B}}(t_1, \ldots, t_n).
$$

Hence

$$
R^{F_k\mathscr{A}}(s_1, \ldots, s_n) \iff (R^{F_k\mathscr{A}}(s_1, \ldots, s_n) \wedge R^{F_k\mathscr{B}}(t_1, \ldots, t_n)) \iff R^{\mathscr{R}}((s_1, t_1), \ldots, (s_n, t_n)),
$$

and thus $p \circ e$ is a strong homomorphism, as required. The verification that q is a pathwise embedding is entirely similar.

Finally, we must show that these morphisms are open. Suppose we are

given a diagram

$$
\begin{array}{ccc}
P & \rightarrowtail & Q \\
\big\downarrow & & \big\downarrow \\
\mathscr{R} & \xrightarrow{\ p\ } & F_k\mathscr{A}
\end{array}
$$

We consider the case where Q extends P by one element; the general case will then follow by induction.

Let the image of the maximum element of P be (s, t), and the image of the maximum element of Q be $s' \succ s$. Since R encodes the plays of a winning strategy for Duplicator, for some $t' \succ t$, we have $(s', t') \in R$, and moreover $(s', t') \in \mathsf{W}(\mathscr{A}, \mathscr{B})$. This implies that the map which extends the embedding $P \rightarrowtail \mathscr{R}$ by sending the maximum element of Q to (s', t') determines an embedding $Q \rightarrowtail \mathscr{R}$, which makes the path-lifting diagram commute.

The fact that q is open is shown by a symmetric argument. $\qquad\square$

For each of our three comonads, we will show that the game $\mathcal{G}_k$ specializes to exactly the expected concrete game: Ehrenfeucht-Fraïssé , pebbling, and modal bisimulation respectively. The connections with the corresponding logical equivalences will follow as immediate corollaries.

10.1 The Ehrenfeucht-Fraïssé comonad

We now show that the generic game $\mathcal{G}_k$, when instantiated to the Ehrenfeucht-Fraïssé adjunction and comonad, yields exactly the standard EF-game.

We begin with two observations we will use.

Lemma 10.2. *Consider $s = [a_1, \ldots, a_i] \in \mathbb{E}_k\mathscr{A}$. If s is non-repeating, then the substructure of $\mathbb{E}_k\mathscr{A}$ determined by the prefixes of s is isomorphic to the substructure of $\mathscr{A}$ determined by $\{a_1, \ldots, a_i\}$.*

Proof. Since the elements of the first set form a chain, this is immediate from the definition of $\mathbb{E}_k\mathscr{A}$, given that the correspondence $[a_1, \ldots, a_j] \mapsto a_j$ is bijective, which holds if s is non-repeating. $\qquad\square$

Lemma 10.3. *The image of an embedding $e : P \rightarrowtail F_k\mathscr{A}$ can only contain non-repeating sequences.*

Proof. Since e is an I-morphism, it must reflect the I relation, while at the same time, it is injective. This implies that $I^{F_k\mathscr{A}}$ must be the identity relation on the image of e, or equivalently, that the image of e must contain only non-repeating sequences. $\qquad\square$

Theorem 10.4. *Given σ-structures $\mathscr{A}$ and $\mathscr{B}$, the $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ game for the Ehrenfeucht-Fraïssé comonad is equivalent to the Ehrenfeucht-Fraïssé game between $\mathscr{A}$ and $\mathscr{B}$.*

Proof. Since $s' \succ s$ in $\mathbb{E}_k\mathscr{A}$ iff $s' = s[a]$ for some $a \in A$, we see from the definition of $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ for the Ehrenfeucht-Fraïssé comonad that it coincides with the Ehrenfeucht-Fraïssé game, provided we can show, for $(s, t) \in \mathbb{E}_k\mathscr{A} \times \mathbb{E}_k\mathscr{B}$, that $(s, t) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$ iff (s, t) satisfies the winning condition for the Ehrenfeucht-Fraïssé game. We recall that this winning condition is that the relation $s_i \mapsto t_i$ is a partial isomorphism between $\mathscr{A}$ and $\mathscr{B}$.

Firstly, suppose that $(s, t) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$, via path embeddings $e_1 : P \rightarrowtail \mathbb{E}_k\mathscr{A}$, $e_2 : P \rightarrowtail \mathbb{E}_k\mathscr{B}$. This implies that the image of e_1, consisting of the prefixes of s, is isomorphic to the image of e_2, consisting of the prefixes of t. Applying lemmas 10.2 and 10.3, we conclude that (s, t) determine a partial isomorphism between $\mathscr{A}$ and $\mathscr{B}$, as required.

For the converse, note that for any (s, t) following a winning strategy for Duplicator in the Ehrenfeucht-Fraïssé game, the correspondence $s_i \mapsto t_i$ must be a partial isomorphism, and in particular a bijection. Hence Duplicator must always give the same response to repetitions of a given move by Spoiler. It follows that the behaviour of the strategy is determined by its restriction to non-repeating sequences, and without loss of generality, we can restrict the strategy to non-repeating sequences (s, t). The prefixes of s form an induced substructure of $\mathbb{E}_k\mathscr{A}$ which, applying lemma 10.2, is isomorphic to the substructure of $\mathbb{E}_k\mathscr{B}$ induced by the prefixes of t. Hence these substructures are the images of path embeddings with a common domain, and $(s, t) \in \mathsf{W}(\mathscr{A}, \mathscr{B})$. $\square$

Theorem 10.5. *For all structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\mathcal{L}_k} \mathscr{B} \iff \mathscr{A} \leftrightarrow_k^{\mathbb{E}} \mathscr{B}$.*

Proof. This is an immediate corollary to Theorems 2.1(1), 10.1, and 10.4. $\square$

Remark If we drop the I-morphism requirement on embeddings, we obtain a weaker notion of equivalence, in which the winning condition is that we have a *partial correspondence* rather than a partial isomorphism. This yields a characterization of elementary equivalence for equality-free logic [11].

10.2 The pebbling comonad

The analysis for the pebbling comonad is largely similar to that for the Ehrenfeucht-Fraïssé comonad.

Firstly, given $s = [(p_1, a_1), \ldots, (p_i, a_i)] \in \mathbb{P}_k\mathscr{A}$, and $p \in \{p_1, \ldots, p_i\}$, we define $\mathsf{last}_p(s) = a$, where $s = s_1[(p, a)]s_2$, and p does not occur in s_2. We say that s is *non-duplicating* if for all $s' \sqsubseteq s$, $\mathsf{last}_p(s') = \mathsf{last}_{p'}(s') \implies p = p'$. Thus we never have two pebbles placed on the same element.

Given $s = [(p_1, a_1), \ldots, (p_i, a_i)] \in \mathbb{P}_k\mathscr{A}$ and $t = [(p_1, b_1), \ldots, (p_i, b_i)] \in \mathbb{P}_k\mathscr{B}$, we define $\gamma(s, t) := \{(\mathsf{last}_p(s), \mathsf{last}_p(t)) \mid p \in \{p_1, \ldots, p_i\}\} \subseteq A \times B$. The prefix of s (respectively t) of length j is denoted s_j (respectively t_j).

The winning condition on plays $(s, t) \in \mathbb{P}_k A \times \mathbb{P}_k B$ for the pebbling game is that for each j, $\gamma(s_j, t_j)$ is a partial isomorphism between $\mathscr{A}$ and $\mathscr{B}$.

The following results play the role for the pebbling comonad of Lemmas 10.2 and 10.3 for the Ehrenfeucht-Fraïssé comonad.

Lemma 10.6. *If s and t are non-duplicating, then the following are equivalent:*

1. *For all j, $\gamma(s_j, t_j)$ is a partial isomorphism from $\mathscr{A}$ to $\mathscr{B}$.*

2. *The substructure P_s of $\mathbb{P}_k\mathscr{A}$ determined by the prefixes of s is isomorphic to the substructure P_t of $\mathbb{P}_k\mathscr{B}$ determined by the prefixes of t.*

Proof. If s and t are non-duplicating, then each $\gamma(s_j, t_j)$ is a partial bijection. Also, the correspondence $\beta : s_j \mapsto t_j$ is a bijection between P_s and P_t. For $s_1, \ldots, s_n \in P_s$, since they are all prefixes of s and hence comparable, $R^{\mathbb{P}_k\mathscr{A}}(s_1, \ldots, s_n)$ iff $R^{\mathscr{A}}(\varepsilon_{\mathscr{A}} s_1, \ldots, \varepsilon_{\mathscr{A}} s_n)$ and $\{\varepsilon_{\mathscr{A}} s_1, \ldots, \varepsilon_{\mathscr{A}} s_n\} \subseteq \mathrm{dom}\,\gamma(s', t')$, where s' is the maximum in the prefix ordering among $s_1, \ldots, s_n$. Similarly for relation instances in P_t. Hence preservation of relation instances by β is equivalent to preservation by $\gamma(s_j, t_j)$ for all j, and similarly for their inverses. $\square$

Lemma 10.7. *The image of an embedding $e : P \rightarrowtail G_k\mathscr{A}$ can only contain non-duplicating sequences.*

Proof. Since e is an I-morphism, it must reflect the I relation, while at the same time, it is injective. This implies that $I^{G_k\mathscr{A}}$ must be the identity relation, or equivalently, that the image of e must contain only non-duplicating sequences. $\square$

Theorem 10.8. *Given σ-structures $\mathscr{A}$ and $\mathscr{B}$, the $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ game for the pebbling comonad is equivalent to the pebbling game between $\mathscr{A}$ and $\mathscr{B}$.*

Proof. The proof proceeds essentially identically to that of Theorem 10.4, except that the appeal to Lemmas 10.2 and 10.3 is replaced by one to Lemmas 10.6 and 10.7. $\square$

Theorem 10.9. *For all structures $\mathscr{A}$ and $\mathscr{B}$: $\mathscr{A} \equiv^{\mathcal{L}^k} \mathscr{B} \iff \mathscr{A} \leftrightarrow_k^{\mathbb{P}} \mathscr{B}$.*

Proof. This is an immediate corollary to Theorems 2.1(2), 10.1, and 10.8. $\square$

10.3 The modal comonad

The key notion of equivalence in modal logic is bisimulation [10, 41]. We shall define the finite approximants to bisimulation [26].[6] Given Kripke structures $\mathscr{A}$ and $\mathscr{B}$, we define a family of relations $\sim_k \subseteq A \times B$: $a \sim_{k+1} b$ iff

1. for all unary P, $P^{\mathscr{A}}(a)$ iff $P^{\mathscr{B}}(b)$

2. for all binary R_α, $R_\alpha^{\mathscr{A}}(a, a')$ implies for some b', $R_\alpha^{\mathscr{B}}(b, b')$ and $a' \sim_k b'$, and $R_\alpha^{\mathscr{B}}(b, b')$ implies for some a', $R_\alpha^{\mathscr{A}}(a, a')$ and $a' \sim_k b'$.

The base case $a \sim_0 b$ holds whenever only the first condition is satisfied.

Plays in the *modal bisimulation game* between pointed structures $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$ are represented by positions $(s, t) \in \mathbb{M}_k(A, a) \times \mathbb{M}_k(B, b)$. A position (s, t) satisfies the winning condition if $s = [a_0, \alpha_1, a_1, \ldots \alpha_i, a_i]$, $t = [b_0, \alpha_1, b_1, \ldots \alpha_i, b_i]$, $a_0 = a$, $b_0 = b$, and for all $0 \leq j \leq i$, and unary predicates P, $P^{\mathscr{A}}(a_j)$ iff $P^{\mathscr{B}}(b_j)$. Note that there is no bijection requirement. This is why I-morphisms are not needed in the modal case.

The initial position is $([a], [b])$. If we have reached position (s, t) after k rounds, then round $k + 1$ proceeds as follows: Spoiler either chooses $s' \succ s$, and Duplicator must respond with $t' \succ t$, producing the new position (s', t'); or Spoiler chooses $t'' \succ t$, and Duplicator must respond with $s'' \succ s$, producing the new position (s'', t''). Duplicator wins the round if they are able to respond, and the new position satisfies the winning condition.

Note that $s' \succ s$ iff $s' = s[\alpha, a']$ for a transition relation R_α and $a' \in A$ such that $R_\alpha^{\mathscr{A}}(a, a')$, where $a = \varepsilon_{\mathscr{A}}(s)$.

The following standard result is essentially immediate from the definitions.

Proposition 10.10. *Given pointed structures $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$, the following are equivalent:*

1. *$a \sim_k b$*

2. *Duplicator has a winning strategy for the k-round modal bisimulation game.*

Theorem 10.11. *Given pointed structures $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$, the $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ game for the modal comonad is equivalent to the modal bisimulation game between $(\mathscr{A}, a)$ and $(\mathscr{B}, b)$.*

Proof. Given the way we have defined the modal bisimulation game, this immediately reduces to verifying that the winning conditions coincide. To

---

[6]Our focus on finite approximants in this paper is for uniformity, and because they are relevant in resource terms. We can extend the comonadic semantics beyond the finite levels. We shall return to this point in the final section.

see this, note that there is a path embedding $e : (R, r) \rightarrowtail \mathbb{M}_k(\mathscr{A}, a)$ with $e(r_i) = s_i$ iff $r = r_0 \xrightarrow{\alpha_1} r_1 \xrightarrow{\alpha_2} \ldots, \xrightarrow{\alpha_i} r_i$, $s_i = [a_0, \alpha_1, a_1, \alpha_2, a_2, \ldots, \alpha_i, a_i]$, with $a = a_0$, and for all unary P, and $0 \leq j \leq i$, $P^Q(r_i) \iff P^\mathscr{A}(a_i)$. Chaining these equivalences for a span of path embeddings shows the equivalence of the winning conditions for $\mathcal{G}_k(\mathscr{A}, \mathscr{B})$ and the modal bisimulation game. $\qquad\square$

Theorem 10.12. *For all pointed structures* $(\mathscr{A}, a)$ *and* $(\mathscr{B}, b)$: $\mathscr{A} \equiv^{\mathbb{M}_k} \mathscr{B} \iff \mathscr{A} \leftrightarrow_k^{\mathbb{M}} \mathscr{B}$.

Proof. This is an immediate corollary to Theorems 2.1(3), 10.1, and 10.11. $\qquad\square$

11 Further Directions

From the categorical perspective, there is considerable additional structure which we have not needed for the results in this paper, but which may be useful for further investigations.

Coequaliser requirements In Moggi's work on computational monads, there is an "equaliser requirement" [36]. The dual version for a comonad (G, ε, δ) is that for every object A, the following diagram is a coequaliser:

$$ G^2 A \xrightarrow[\varepsilon_{GA}]{\overset{G\varepsilon_\mathscr{A}}{\longrightarrow}} GA \xrightarrow{\varepsilon_\mathscr{A}} A $$

This says in particular that the counit is a regular epi, and hence GA "covers" A in a strong sense.

This coequaliser requirement holds for all our comonads. For $\mathbb{E}_k$, this is basically the observation that, given a sequence of sequences $[s_1, \ldots, s_j]$, we have $\varepsilon[\varepsilon s_1, \ldots, \varepsilon s_j] = \varepsilon s_j$. The other cases are similar.

Indexed and graded structure Our comonads $\mathbb{E}_k$, $\mathbb{P}_k$, $\mathbb{M}_k$ are not merely discretely indexed by the resource parameter. In each case, there is a functor $(\mathbb{Z}^+, \leq) \to \mathsf{Comon}(\mathcal{R}(\sigma))$ from the poset category of the positive integers to the category of comonads on $\mathcal{R}(\sigma)$. Thus if $k \leq l$ there is a natural transformation with components $i_A^{k,l} : \mathbb{E}_k\mathscr{A} \to \mathbb{E}_l\mathscr{A}$, which preserves the counit and comultiplication; and similarly for the other comonads. Concretely, this is just including the plays of up to k rounds in the plays of up to l rounds, $k \leq l$.

Another way of parameterizing comonads by resource information is grading [21]. Recall that comonads on $\mathcal{C}$ are exactly the comonoids in the strict monoidal category $([\mathcal{C}, \mathcal{C}], \circ, I)$ of endofunctors on $\mathcal{C}$ [32]. Generalizing this description, a graded comonad is an oplax monoidal functor $G : (M, \cdot, 1) \to ([\mathcal{C}, \mathcal{C}], \circ, I)$ from a monoid of grades into this endofunctor category. This means that for each $m \in M$, there is an endofunctor G_m,

there is a graded counit natural transformation $\varepsilon : G_1 \Rightarrow I$, and for all $m, m' \in M$, there is a graded comultiplication $\delta^{m,m'} : G_{m \cdot m'} \Rightarrow G_m G_{m'}$.

The two notions can obviously be combined. We can see our comonads as (trivially) graded, by viewing them as oplax monoidal functors $(\mathbb{Z}_\infty^+, \leq, \min, \infty) \to ([\mathcal{C}, \mathcal{C}], \circ, I)$. Given $k \leq l$, we have e.g. $\mathbb{E}_k \Rightarrow \mathbb{E}_k \mathbb{E}_k \Rightarrow \mathbb{E}_k \mathbb{E}_l$. The unit is interpreted using the ω-colimit $\mathbb{E}_\omega$, as described in the following paragraph.

The question is whether there are interesting graded structures which arise naturally in considering richer logical and computational situations in our setting.

Colimits and infinite behaviour In this paper, we have dealt exclusively with finite resource levels. However, there is an elegant means of passing to infinite levels. We shall illustrate this with the modal comonad. Using the inclusion morphisms described in the previous discussion of indexed structure, for each pointed structure $(\mathscr{A}, a)$ we have a diagram

$$\mathbb{M}_1(\mathscr{A}, a) \to \mathbb{M}_2(\mathscr{A}, a) \to \cdots \to \mathbb{M}_k(\mathscr{A}, a) \to \cdots$$

By taking the colimits of these diagrams, we obtain a comonad $\mathbb{M}_\omega$, which corresponds to the usual unfolding of a Kripke structure to all finite levels. This will correspond to the bisimulation approximant $\sim_\omega$, which coincides with bisimulation itself on image-finite structures [26]. Transfinite extensions are also possible. Similar constructions can be applied to the other comonads. This provides a basis for lifting the comonadic analysis to the level of infinite models.

Relations between fragments and parameters We can define morphisms between the different comonads we have discussed, which yield proofs about the relationships between the logical fragments they characterize. This categorical perspective avoids the cumbersome syntactic translations in the standard proofs of these results. For illustration, there is a comonad morphism $t : \mathbb{E}_k \Rightarrow \mathbb{P}_k$ with components $t_A : \mathbb{E}_k \mathscr{A} \to \mathbb{P}_k \mathscr{A}$ given by $[a_1, \ldots, a_j] \mapsto [(1, a_1), \ldots, (j, a_j)]$. Together with theorems 5.4 and 5.3, this shows that $\exists \mathcal{L}_k \subseteq \exists \mathcal{L}^k$ and $\mathcal{L}_k(\#) \subseteq \mathcal{L}^k(\#)$. Moreover, composing t with a coalgebra $\mathscr{A} \to \mathbb{E}_k \mathscr{A}$ yields a coalgebra $A \to \mathbb{P}_k \mathscr{A}$, demonstrating that $\mathsf{tw}(\mathscr{A}) + 1 \leq \mathsf{td}(\mathscr{A})$.

Another morphism $\mathbb{M}_\omega \Rightarrow \mathbb{P}_2$ shows that modal logic can be embedded into 2-variable logic. More precisely, we can lift $\mathbb{P}_2$ to $\mathcal{R}_\star(\sigma)$ by defining the distinguished element of $\mathbb{P}_2(\mathscr{A}, a)$ to be $[(1, a)]$. The component of the morphism at $(\mathscr{A}, a)$ sends $[a_0, \alpha_1, a_1, \ldots, \alpha_j, a_j]$ to $[(1, a_0), (2, a_1), (1, a_2), \ldots ((j \bmod 2) + 1, a_j)]$. This captures the "hand-over-hand" reuse of variables in a syntax-free fashion.

Concluding remarks

Our comonadic constructions for the three major forms of model comparison games show a striking unity, on the one hand, but also some very interesting

differences. For the latter, we note the different forms of logical "deception" associated with each comonad, the finite character of $\mathbb{E}_k$ and $\mathbb{M}_k$ and the non-finite character of $\mathbb{P}_k$, and the different combinatorial parameters which arise in each case.

One clear direction for future work is to gain a deeper understanding of what makes these constructions work. Another is to understand how widely the comonadic analysis of resources can be applied. We are currently investigating the guarded fragment [9, 23]; other natural candidates include existential second-order logic, and branching quantifiers and dependence logic [46].

Since comonads arise naturally in type theory and functional programming [45, 38], can we connect the study of finite model theory made here with a suitable type theory? Can this lead, via the Curry-Howard correspondence, to the systematic derivation of some significant meta-algorithms, such as decision procedures for guarded logics based on the tree model property [22], or algorithmic metatheorems such as Courcelle's theorem [15]?

Another intriguing direction is to connect these ideas with the graded quantum monad studied in [2], which provides a basis for the study of quantum advantage in $\mathcal{R}(\sigma)$. This may lead to a form of quantum finite model theory.

Current and ongoing work

The first author of the present paper and Anuj Dawar have recently begun a U.K. EPSRC-funded joint project on "Resources and Co-Resources: a junction between categorical semantics and descriptive complexity" to pursue the ideas introduced in [3] and the present paper. Project participants include the second author of the present paper, Tom Paine, Adam Ó Conghaile, Daniel Marsden, Tomáš Jakl and Luca Reggio. A categorical analysis of Rossman's Equirank HPT Theorem appears in [1], and a significant refinement of this result, involving a combination of the pebbling and Ehrenfeucht-Fraïssé comonads, is presented in [39]. Current work in progress includes comonadic treatments of guarded fragments [4] and of generalized quantifiers [13], and coalgebraic characterizations of further combinatorial invariants such as clique-width.

References

[1] Samson Abramsky. Whither semantics? *Theoretical Computer Science*, 807:3–14, 2020.

[2] Samson Abramsky, Rui Soares Barbosa, Nadish de Silva, and Octavio Zapata. The quantum monad on relational structures. In *Proceed-*

ings of 42nd International Symposium on Mathematical Foundations of Computer Science, MFCS 2017, page 35:1–35:19. Springer, 2018.

[3] Samson Abramsky, Anuj Dawar, and Pengming Wang. The pebbling comonad in finite model theory. In *Logic in Computer Science (LICS), 2017 32nd Annual ACM/IEEE Symposium on*, pages 1–12. IEEE, 2017.

[4] Samson Abramsky and Dan Marsden. Comonadic semantics for guarded fragments. *arXiv preprint arXiv:2008.11094*, 2020.

[5] Samson Abramsky and Nihil Shah. Relating structure and power: Comonadic semantics for computational resources. In *27th EACSL Annual Conference on Computer Science Logic, CSL 2018, September 4-7, 2018, Birmingham, UK*, pages 2:1–2:17, 2018. `doi:10.4230/LIPIcs.CSL.2018.2`.

[6] Luca Aceto, Anna Ingolfsdottir, and Joshua Sack. Resource bisimilarity and graded bisimilarity coincide. *Information Processing Letters*, 111(2):68–76, 2010.

[7] Danel Ahman, James Chapman, and Tarmo Uustalu. When is a container a comonad? In *International Conference on Foundations of Software Science and Computational Structures*, pages 74–88. Springer, 2012.

[8] Thorsten Altenkirch, James Chapman, and Tarmo Uustalu. Monads need not be endofunctors. In *International Conference on Foundations of Software Science and Computational Structures*, pages 297–311. Springer, 2010.

[9] Hajnal Andréka, István Németi, and Johan van Benthem. Modal languages and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27(3):217–274, 1998.

[10] Patrick Blackburn, Maarten De Rijke, and Yde Venema. *Modal Logic*, volume 53. Cambridge University Press, 2002.

[11] E. Casanovas, P. Dellunde, and R. Jansana. On Elementary Equivalence for Equality-free Logic. *Notre Dame Journal of Formal Logic*, 37(3):506–522, 1996. URL: `http://projecteuclid.org:80/Dienst/getRecord?id=euclid.ndjfl/1039886524/`, `doi:10.1305/ndjfl/1039886524`.

[12] Ashok K Chandra and Philip M Merlin. Optimal implementation of conjunctive queries in relational data bases. In *Proceedings of the Ninth Annual ACM Symposium on Theory of Computing*, pages 77–90. ACM, 1977.

[13] Adam Ó Conghaile and Anuj Dawar. Game comonads & generalised quantifiers. *arXiv preprint arXiv:2006.16039*, 2020.

[14] Flavio Corradini, Rocco De Nicola, and Anna Labella. Graded modalities and resource bisimulation. In *International Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 381–393. Springer, 1999.

[15] Bruno Courcelle. The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. *Information and computation*, 85(1):12–75, 1990.

[16] Víctor Dalmau, Phokion G Kolaitis, and Moshe Y Vardi. Constraint satisfaction, bounded treewidth, and finite-variable logics. In *International Conference on Principles and Practice of Constraint Programming*, pages 310–326. Springer, 2002.

[17] M. de Rijke. A Note on Graded Modal Logic. *Studia Logica*, 64(2):271–283, 2000.

[18] Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite model theory*. Springer Science & Business Media, 2005.

[19] Andrzej Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Math*, 49(129-141):13, 1961.

[20] Roland Fraïssé. Sur quelques classifications des systèmes de relations. In *Publications Scientifiques, Série A*, pages 35–182. Université d'Alger, 1954.

[21] Marco Gaboardi, Shin-ya Katsumata, Dominic Orchard, Flavien Breuvart, and Tarmo Uustalu. Combining effects and coeffects via grading. *ACM SIGPLAN Notices*, 51(9):476–489, 2016.

[22] Erich Grädel. Decision procedures for guarded logics. In *International Conference on Automated Deduction*, pages 31–51. Springer, 1999.

[23] Erich Grädel and Martin Otto. The freedoms of (guarded) bisimulation. In *Johan van Benthem on Logic and Information Dynamics*, pages 3–31. Springer, 2014.

[24] Pavol Hell and Jaroslav Něsetřil. *Graphs and homomorphisms*. Oxford University Press, 2004.

[25] Lauri Hella. Logical hierarchies in PTIME. *Information and Computation*, 121:1–19, 1996.

[26] Matthew Hennessy and Robin Milner. On observing nondeterminism and concurrency. In *International Colloquium on Automata, Languages, and Programming*, pages 299–309. Springer, 1980.

[27] André Joyal, Mogens Nielson, and Glynn Winskel. Bisimulation and open maps. In *[1993] Proceedings Eighth Annual IEEE Symposium on Logic in Computer Science*, pages 418–427. IEEE, 1993.

[28] Ton Kloks. *Treewidth: computations and approximations*, volume 842. Springer Science & Business Media, 1994.

[29] Phokion G Kolaitis and Moshe Y Vardi. On the expressive power of Datalog: tools and a case study. In *Proceedings of the ninth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, pages 61–71. ACM, 1990.

[30] Phokion G Kolaitis and Moshe Y Vardi. Infinitary logics and 0–1 laws. *Information and Computation*, 98(2):258–294, 1992.

[31] Leonid Libkin. *Elements of Finite Model Theory (Texts in Theoretical Computer Science. An EATCS Series)*. Springer, 2004.

[32] Saunders Mac Lane. *Categories for the working mathematician*, volume 5. Springer Science & Business Media, 2013.

[33] Ernest G Manes. *Algebraic Theories*, volume 26. Springer Science & Business Media, 2012.

[34] Robin Milner. *A calculus of communicating systems*. Number 92 in Lecture Notes in Comput. Science. Springer-Verlag, 1980.

[35] Robin Milner. *Communication and concurrency*, volume 84. Prentice Hall New York etc., 1989.

[36] Eugenio Moggi. Notions of computation and monads. *Information and computation*, 93(1):55–92, 1991.

[37] Jaroslav Nešetřil and Patrice Ossona De Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European Journal of Combinatorics*, 27(6):1022–1041, 2006.

[38] Dominic Orchard. Programming contextual computations. Technical Report UCAM-CL-TR-854, University of Cambridge, 2014.

[39] Tom Paine. A pebbling comonad for finite rank and variable logic, and an application to the equirank-variable homomorphism preservation theorem. In *Proceedings of MFPS*, 2020.

[40] Benjamin Rossman. Homomorphism Preservation Theorems. *JACM*, 55(3):15, 2008. URL: http://www.math.toronto.edu/rossman/hpt-jacm.pdf.

[41] Davide Sangiorgi. On the origins of bisimulation and coinduction. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 31(4):15, 2009.

[42] Saharon Shelah. Every two elementarily equivalent models have isomorphic ultrapowers. *Israel Journal of Mathematics*, 10(2):224–233, 1971.

[43] JD Ullman. *Database and Knowledge-Based Systems, Vols. I and II.* Computer Science Press, Rockville, MD, 1989.

[44] Friedrich Ulmer. Properties of dense and relative adjoint functors. *Journal of algebra*, 8(1):77–95, 1968.

[45] Tarmo Uustalu and Varmo Vene. Comonadic notions of computation. *Electronic Notes in Theoretical Computer Science*, 203(5):263–284, 2008.

[46] Jouko Väänänen. *Dependence logic: A new approach to independence friendly logic*, volume 70. Cambridge University Press, 2007.

[47] Glynn Winskel and Mogens Nielsen. Models for Concurrency. Handbook of Logic in Computer Science, Vol. IV, ed. Abramsky, Gabbay and Maibaum. *Oxford University Press*, 1(995):1–148, 1995.