A model-theoretic approach to descriptive general frames: the van Benthem characterisation theorem

Nick Bezhanishvili and Tim Henke

Abstract

The celebrated van Benthem characterisation theorem states that on Kripke structures modal logic is the bisimulation-invariant fragment of first-order logic. In this paper we prove an analogue of the van Benthem characterisation theorem for models based on descriptive general frames. This is an important class of general frames for which every modal logic is complete. These frames can be represented as Stone spaces equipped with a "continuous" binary relation. The proof of our theorem generalises Rosen's proof of the van Benthem theorem for finite frames and uses as an essential technique a new notion of descriptive unravelling. We also develop a basic model theory for descriptive general frames and show that in many ways it behaves like the model theory of finite structures. In particular, we prove the failure of the Compactness theorem, of the Beth definability theorem, of the Craig interpolation theorem, and of the Upward Löwenheim-Skolem theorem.[1]

1 Introduction

Kripke models are relational structures which provide standard models for modal logic. Bisimulations are relations on Kripke models that are indistinguishable by modal logic, in the sense that the truth of modal formulae is preserved under these relations. The van Benthem characterisation theorem states that in fact, any first-order formula that is invariant under bisimulations must be equivalent to a modal formula [1, 4]. This is often formulated more succinctly by saying that modal logic is the bisimulation-invariant fragment of first-order logic. The theorem has inspired many generalisations and alternative versions, including a similar characterisation for intuitionistic logic [19, 20], neighbourhood models [12], and numerous coalgebraic generalisations [22]. Another notable result is the Janin-Walukiewicz theorem [15], showing that the modal μ-calculus is the bisimulation-invariant fragment of monadic second-order logic.

A result that is of particular importance to this paper was given by Rosen in [21], showing that over finite models, too, modal logic is the bisimulation-invariant fragment of first-order logic. This is particularly remarkable because

[1] This paper is based on [14].

the Compactness theorem of first-order logic features prominently in the proof of the original van Benthem characterisation theorem, while the class of finite models crucially lacks the compactness property. The result for finite models has in turn been generalised to multiple other classes lacking the compactness property, including the classes of rooted finite models, rooted transitive models, and well-founded transitive models [6].

The main result of this paper is the van Benthem characterisation theorem for the class of models over descriptive general frames. These frames can be represented topologically as totally separated, compact topological spaces, commonly known as Stone spaces, with a continuous relation, whose models have valuations restricted to clopen sets.

Descriptive general frames form an important class of frames with respect to which, unlike standard Kripke semantics, every modal logic is complete. This is due to the Jónsson-Tarski duality between descriptive general frames and modal algebras [4, 5, 16],

Descriptive general frames are also a natural generalisation of finite frames. Topological compactness is a common generalisation of finiteness and the classes of finite and descriptive general frames are both closed under many operations such as finite disjoint unions and p-morphic images, they both are not closed under infinite disjoint unions and they both have the Hennessy-Milner property (see [4, Theorem 2.24] and [2, Corollary 3.10]). Moreover, we show that they have a similar model theory. In particular, we will prove that the Compactness theorem, the Beth definability theorem, the Craig interpolation theorem, and the Upward Löwenheim-Skolem theorem all fail on descriptive general frames.

For the proof of the van Benthem theorem our central tool is, what we call, descriptive unravelling. This is some kind of a descriptive completion of the unravelling tree commonly seen and used in modal logic. Roughly speaking, we add new points to the standard unravelling at "infinite distance" from the points of unravelling, yet guaranteeing that the new model is descriptive. Proving a number of important properties of this construction then allows for its use in the main proof through an Ehrenfeucht-Fraïssé argument.

Finally, we point out that descriptive general frames can be viewed as coalgebras for the Vietoris functor on the category of Stone spaces [18], just like Kripke frames are coalgebras for the powerset functor and finite Kripke frames are coalgebras for the powerset functor restricted to finite sets [23]. While an analogue of the van Benthem theorem for coalgebraic logics given by Set-functors has been proved in [22], as far as we are aware, the result in this paper is the first van Benthem-style characterisation theorem for coalgebras over a category of topological spaces. This may pave the way for generalisations of the theorem for coalgebras of other Vietoris-like functors. These and other potential generalisations are briefly discussed in the conclusions.

The paper is organised as follows: In Section 2, descriptive general frames and their models are introduced and discussed. Algebraic duality for general frames is presented and a descriptive completion is introduced. Additionally, Vietoris bisimulations for models based on general frames are defined and finite approximations to bisimulations are briefly discussed. Section 3 focuses on the model theory of descriptive models and proves the failure of a number of important results from classical model theory such as the Compactness theorem for first-order logic and the Upward Löwenheim-Skolem theorem, on the class of descriptive general models. Inspired by the model-theoretic similarities between the classes of finite and descriptive models, Section 4 uses the approach from [21] to prove our main result, the van Benthem characterisation theorem for descriptive models. An unravelling constructionm which is the central tool in the proof is presented together with a duplication procedure. Some of their properties are stated and proven, and with those results, the main theorem is shown through an Ehrenfeucht-Fraïssé argument. The final section contains a brief summary of the paper and points to a number of possible new research directions.

2 Preliminaries

In this section we briefly discuss some of the basic definitions and facts that will be used throughout the paper. We use [4, 5, 17] as our main references for the basic theory of modal logic. We assume the reader's familiarity with the basic concepts of modal logic.

2.1 Descriptive frames

Let (W, R) be a Kripke frame, i.e., W is a set and $R \subseteq W \times W$ is a binary relation on W. For each $x \in W$ and $U \subseteq W$ we let $R[x] = \{y \in W : xRy\}$, $R^{-1}[x] = \{y \in W : yRx\}$, and $\langle R \rangle U = \{x \in W : R[x] \cap U \neq \emptyset\}$.

We next recall the general frame semantics for modal logic.

Definition 2.1. A triplet $\mathfrak{g} = (W, R, A)$ is a *general frame* if (W, R) is a Kripke frame and A is a field of sets[2] over W that is closed under the operation $\langle R \rangle$ (i.e., for every $a \in A$ also $\langle R \rangle a \in A$). The underlying frame (W, R) of $\mathfrak{g}$ will be denoted by $\mathfrak{g}_{\#}$.

A quadruplet $\mathfrak{m} = (W, R, A, V)$ is a *general model* based on $\mathfrak{g} = (W, R, A)$ over a set of propositional variables P if $V : \mathsf{P} \to A$ is a function. The definition of $[\![\cdot]\!]^{\mathfrak{m}}$ is the same for general models as it is for Kripke models. ◀

This definition is motivated by the fact that now, the truth set $[\![\varphi]\!]^{\mathfrak{m}}$ belongs to A for any general model $\mathfrak{m}$ and formula φ.

General frames also have a natural structure of a topological space, with A acting as a basis of clopens on a universe W. This opens the subject for topological analysis, which has led in the past to many crucial insights into these structures.

Some specific types of general frames are of special importance. The following classes of important general frames will be relevant to this paper.

[2]That is, a non-empty collection of sets that is closed under binary union and complementation.

Definition 2.2. Let $\mathfrak{g} = (W, R, A)$ be a general frame. Then

- $\mathfrak{g}$ is called *differentiated* if for all distinct $w, v \in W$ there exists an $a \in A$ such that $v \in a$ but $w \notin a$;

- $\mathfrak{g}$ is *tight*[3] if for all $w, v \in W$ with $(w, v) \notin R$ there exists an $a \in A$ such that $v \in a$ and $w \notin \langle R \rangle a$;

- $\mathfrak{g}$ is called *compact* if all collections $\mathcal{A} \subseteq A$ with empty intersection have a finite subcollection $\mathcal{A}_0 \subseteq \mathcal{A}$ with empty intersection;

- $\mathfrak{g}$ is called *image-compact* if for any point $w \in W$, a collection $\mathcal{A} \subseteq A$ whose intersection is disjoint from $R[w]$, the set of successors of w, has a finite subcollection $\mathcal{A}_0 \subseteq \mathcal{A}$ whose intersection is also disjoint from $R[w]$;

- $\mathfrak{g}$ is called *descriptive* if it is differentiated, tight, and compact. The class of models based on descriptive frames will be denoted by $\mathscr{D}$.

Models based on such a general frame are called *differentiated, tight, compact,* or *descriptive models.* In the presence of these modifiers, the adjective "general" will usually be omitted from both the frames and the models. ◀

The term compactness is used because it corresponds to topological compactness. Differentiatedness is equivalent to the topological space being totally separated (i.e., every two distinct points being separated by a clopen set). Topological spaces that are compact and totally separated are called *Stone spaces.* Descriptive frames are then pairs (W, R) where W is a Stone space and R is such that

- $R[x]$ is closed for every $x \in W$;

- For every clopen set $U \subseteq W$ the set $\langle R \rangle U$ is also clopen.

We refer to [4, 17] for a detailed discussion on the connection of these two approaches to descriptive frames. From now on we will use the topological concept interchangeably with the subclass of general frames.

2.2 Duality

One of the many reasons descriptive frames have proven a valuable tool, is their role in Jónsson-Tarski duality for modal algebras. This duality builds on the celebrated Stone duality between the category of Boolean algebras and Boolean algebra homomorphisms is dually equivalent to the category of Stone spaces and continuous maps.

Definition 2.3. A *modal algebra* is a pair $\mathcal{M} = (\mathcal{B}, \Diamond)$ where $\mathcal{B} = (B, \vee, \wedge, \neg, 0, 1)$ is a Boolean algebra, and $\Diamond : B \to B$ is an operator satisfying $\Diamond 0 = 0$ and $\Diamond(a \vee b) = \Diamond a \vee \Diamond b$.

[3] The tightness condition can also be considered a continuity condition for the relation with respect to the topology, see also [14, Definition 2.71, Proposition 2.73].

The category **MA** has as objects all modal algebras and as morphisms the Boolean morphisms f such that $f(\Diamond a) = \Diamond f(a)$ for all a. ◀

Definition 2.4. The category **DF** is the category with all descriptive frames as objects, and as morphisms the continuous bounded morphisms.[4] ◀

The functors that establish the category duality will be referred to as $(-)_*$ and $(-)^*$. We will briefly recall their definition.

Let $\mathcal{M} = (\mathcal{B}, \Diamond)$ be a modal algebra. Then let $\mathcal{M}_* = (W, R, A)$ be the descriptive frame (see [4, Theorem 5.76]) with

$$\begin{aligned} W &= \mathrm{Uf}\ \mathcal{B} && \text{the collection of ultrafilters on } \mathcal{B}; \\ A &= \left\{\widehat{b} \mid b \in \mathcal{B}\right\} && \text{where } \widehat{b} := \{F \in W \mid b \in F\}; \\ FRF' &:\Longleftrightarrow b \in F' \text{ implies } \Diamond b \in F. \end{aligned}$$

Moreover, if $f : \mathcal{M} \to \mathcal{M}'$ is a morphism of modal algebras, let then $f_* : \mathcal{M}'_* \to \mathcal{M}_*$ be the map given by $f_*(F') = f^{-1}[F']$ for each ultrafilter $F' \in \mathcal{M}'_*$. In [4, Proposition 5.80] this is shown to be a continuous bounded morphism.

Let $\mathfrak{g} = (W, R, A)$ be a descriptive frame. Then the associated modal algebra is $\mathfrak{g}^* = (A, \langle R \rangle)$. Continuous bounded morphisms $f : \mathfrak{g} \to \mathfrak{g}'$ are mapped to the map $f^* : (\mathfrak{g}')^* \to \mathfrak{g}^*$ of modal algebras given by $f^*(a) = f^{-1}[a]$ for each $a \in A$.

Theorem 2.5 (Jónsson-Tarski duality [4, 5, 16, 17])**.** *The functors $(-)_*$ and $(-)^*$ provide a dual category equivalence between* **MA** *and* **DF**.

It is important to note the use of the $(-)^*$-functor is not restricted to descriptive frames only. For the entire class of general frames, one may define the exact same operation. Combining it with the other functor results in an operation that turns every general frame into a descriptive frame. This may be called the descriptive completion of the general frame.

Definition 2.6. Let $\mathfrak{g} = (W, R, A)$ be a general frame. Then its *descriptive completion* is $(\mathfrak{g}^*)_* =: (W_*, R_*, \widehat{A})$, where $\widehat{A} = \{\widehat{a} \mid a \in A\}$. If $\mathfrak{m} = (\mathfrak{g}, V)$ is a general model, there is an induced descriptive model $\mathfrak{m}_* := ((\mathfrak{g}^*)_*, V_*)$ where V_* is the unique valuation such that if originally $V(p) = a$, then $V_*(p) = \widehat{a}$. ◀

The name "descriptive completion" is justified by the following theorem.

Theorem 2.7. ([4, Theorem 5.76]) *Let $\mathfrak{g}$ be a general frame. Then*

- $(\mathfrak{g}^*)_*$ *is a descriptive frame,*
- $\mathfrak{g} \cong (\mathfrak{g}^*)_*$ *if and only if $\mathfrak{g}$ is descriptive.*

[4] *A bounded morphism* between Kripke frames $\mathfrak{F} = (W, R)$ and $\mathfrak{F}' = (W', R')$ is a map $f : W \to W'$ with the property that if wRv then $f(w)Rf(v)$ and if $f(w)R'v'$ then there is a $v \in W$ such that wRv and $f(v) = v'$. A map is continuous if the inverse image of any open set is open.

In the special case that $\mathfrak{g}$ is a Kripke frame, meaning that $A = \mathscr{P}(W)$ is the powerset, then $(\mathfrak{g}^*)_*$ is known as the *ultrafilter extension* of $\mathfrak{g}$, see also [4, Section 2.5].

This completion construction will be used in Section 4 to develop the central tool in the proof for a van Benthem-type result.

2.3 Kripke and Vietoris bisimulations

An important notion in modal logic that is central to main theorem of this thesis is the concept of bisimulations. Intuitively, if two points w and v in models $\mathfrak{M}$ and $\mathfrak{N}$ respectively are bisimilar, they are impossible to distinguish by walking from w and v through the models and looking only at the propositional variables that are true at the points reached. In this section we will recall the notions of Kripke and Vietoris bisimulations for descriptive models.

Definition 2.8. Let $\mathfrak{M}$ and $\mathfrak{M}'$ be (general) models on frames (W, R) and (W', R') with valuations V and V'. A relation $Z \subseteq W \times W'$ is called a (*Kripke*) *bisimulation* if for all $(w, w') \in Z$ the following three conditions hold:

(prop) If $p \in \mathsf{P}$ then $w \in V(p) \iff w' \in V'(p)$;

(forth) If $v \in R[w]$ then there exists a $v' \in R'[w']$ such that $(v, v') \in Z$;

(back) If $v' \in R'[w']$ then there exists a $v \in R[w]$ such that $(v, v') \in Z$.

If two points w and v in models $\mathfrak{M}$ and $\mathfrak{N}$ are linked by a bisimulation, then we say that w and v are *bisimilar*. A model $\mathfrak{M}$ with a fixed element w will be called a *pointed model*. If two points w and v in models $\mathfrak{M}$ and $\mathfrak{N}$ respectively are bisimilar, then we say that pointed models $\mathfrak{M}, w$ and $\mathfrak{N}, v$ are bisimilar. ◀

As will be shown in Theorem 2.15, two points linked by a bisimulation satisfy the same modal formulae.

Definition 2.9 ([2, 8])**.** Let $\mathfrak{m}$ and $\mathfrak{m}'$ be general models and Z a Kripke bisimulation between them. If Z is closed in the product topology of the two associated topological spaces, then it is called a *Vietoris bisimulation.* ◀

If two pointed models are linked by a Kripke bisimulation this is denoted by

$$\mathfrak{m}, w \leftrightarroweq \mathfrak{m}', w'$$

and if they are linked by a Vietoris bisimulation this will be written as

$$\mathfrak{m}, w \leftrightharpoons \mathfrak{m}', w'.$$

Remark 2.10. This definition is motivated by a coalgebraic perspective on descriptive frames, in which it is the coalgebraic definition of a bisimulation on descriptive frames as coalgebras for the Vietoris functor. However, on descriptive frames this is equivalent to the definition provided above. A detailed exploration of the coalgebraic notion of descriptive frames can be found in [2, 18, 24] and an extensive treatment of Vietoris bisimulations, including this equivalence, can be found in [2, 8]. ◀

Theorem 2.11 ([2, Theorem 5.2])**.** *If $\mathfrak{m}, w$ and $\mathfrak{n}, v$ are two pointed image-compact general models, then the following are equivalent:*

1. $\mathfrak{m}, w \leftrightarroweq \mathfrak{n}, v$;
2. $\mathfrak{m}, w \leftrightharpoons \mathfrak{n}, v$.

2.4 Finite bisimulations

Crucial for this paper will be the notion of finite (approximations to) bisimulations. Like the finitary modal language, the finite number of steps involved makes them much easier to deal with.

Definition 2.12. Let $k \in \mathbb{N}$ be a natural number, P a set of propositional variables and $\mathfrak{M}$ and $\mathfrak{M}'$ be two (general) models with frames (W, R) and (W', R') and valuations V and V' over P respectively. Then a k-bisimulation over P is a $\subseteq$-decreasing $(k+1)$-sequence $(Z_\ell)_{0 \le \ell \le k}$ of relations $Z_\ell \subseteq W \times W'$ such that for all natural numbers $\ell \le k$ and $(w, w') \in Z_\ell$:

(prop) If $p \in \mathsf{P}$ then $w \in V(p)$ if and only if $w' \in V'(p)$;

(forth) For all non-negative $m < \ell$, if $v \in R[w]$ then there exists a $v' \in R'[w']$ such that the pair (v, v') is an element of Z_m;

(back) For all non-negative $m < \ell$, if $v' \in R'[w']$ then there exists a $v \in R[w]$ such that the pair (v, v') is an element of Z_m.

If there exists an k-bisimulation $(Z_\ell)_{0 \le \ell \le k}$ with $(w, v) \in Z_k$, this is denoted by $\mathfrak{M}, w \leftrightarroweq_k \mathfrak{N}, v$. ◀

One could also define finite Vietoris bisimulations when all Z_ℓ are closed, but this notion will not be useful for this paper.

Remark 2.13. If Z is a Kripke bisimulation, then for any $k \in \mathbb{N}$ the sequence $(Z)_{0 \le \ell \le k}$ is a k-bisimulation. As such, $\mathfrak{M}, w \leftrightarroweq \mathfrak{N}, v$ implies $\mathfrak{M}, w \leftrightarroweq_k \mathfrak{N}, v$ for any k. ◀

Lemma 2.14. Let $k \in \mathbb{N}$ and $\mathfrak{M}, w$ and $\mathfrak{N}, v$ be two pointed (general) models and P a set of propositional variables. Then

$$\text{If } \mathfrak{M}, w \leftrightarroweq_k \mathfrak{N}, v \text{ over } \mathsf{P} \text{ then } \mathfrak{M}, w \leftrightsquigarrow_k \mathfrak{N}, v \text{ over } \mathsf{P},$$

where $\leftrightsquigarrow_k$ denotes agreement on all modal formulae on modal depth $\le k$. Moreover, if P is finite, then the converse implication holds as well.

Proof. The proof for Kripke models can be found in e.g. [4, Proposition 2.31]. As general models satisfy the same formulae as their associated Kripke models[5] and the definition of k-bisimulations references only the Kripke model structure, this immediately extends the result to general models. □

[5]That is, a general model (W, R, A, V) has an associated Kripke model (W, R, V), which satisfies the same formulae in the same points as the collection of admissible sets, whose only semantic function is to restrict the choice of valuation, is not relevant when the valuation has already been chosen.

Theorem 2.15. *Any two (Vietoris-)bisimilar pointed (general) models satisfy the same formulae.*

Proof. This follows immediately from Lemma 2.14 and Remark 2.13 for the Kripke case, and additionally Theorem 2.11 for the Vietoris case. □

An important note to make is that these are equivalence relations, in particular transitive.

Lemma 2.16. Assume that $\mathfrak{M}_0, w_0 \underline{\leftrightarrow}_\ell \mathfrak{M}_1, w_1$ and $\mathfrak{M}_1, w_1 \underline{\leftrightarrow}_k \mathfrak{M}_2, w_2$ for $\ell \leq k$. Then $\mathfrak{M}_0, w_0 \underline{\leftrightarrow}_\ell \mathfrak{M}_2, w_2$. In particular, if $\mathfrak{M}_0, w_0 \underline{\leftrightarrow} \mathfrak{M}_1, w_1$ and $\mathfrak{M}_0, w_0 \underline{\leftrightarrow}_\ell \mathfrak{N}, v$, then $\mathfrak{M}_1, w_1 \underline{\leftrightarrow}_\ell \mathfrak{N}, v$.

Proof. Let $(Z_m)_{0 \leq m \leq \ell}$ be an ℓ-bisimulation between $\mathfrak{M}_0, w_0$ and $\mathfrak{M}_1, w_1$ and let $(\widetilde{Z}_n)_{0 \leq n \leq k}$ be a k-bisimulation between $\mathfrak{M}_1, w_1$ and $\mathfrak{M}_2, w_2$. Then consider $(Z_m; \widetilde{Z}_m)_{0 \leq m \leq \ell}$ as an ℓ-bisimulation between $\mathfrak{M}_0, w_0$ and $\mathfrak{M}_2, w_2$, where ; denotes the composition of relations. Suppose that $(v_0, v_2) \in Z_m; \widetilde{Z}_m$ for some $m \leq \ell$. Then there exists a v_1 such that $v_0 Z_m v_1$ and $v_1 \widetilde{Z}_m v_2$. Consequently, v_0 satisfies the same propositional variables as v_1, which in turn satisfies the same propositional variables as v_2, verifying that v_0 and v_2 satisfy the same propositional variables.

For the forth condition, if v_0 has a successor x_0, then there is a successor x_1 of v_1 such that $x_0 Z_n x_1$ for all $n < m$. Then from $v_1 \widetilde{Z}_m v_2$ it follows that there is a successor x_2 of v_2 such that $x_1 \widetilde{Z}_n x_2$ for all $n < m$. Therefore $(x_0, x_2) \in (Z_n; \widetilde{Z}_n)$ for all $n < m$. The back condition is identical. The final observation then follows from Remark 2.13. □

3 Model-theoretic failures on descriptive models

In this section we will study the basic model-theoretic properties of descriptive frames. We will show that, similarly to finite frames, many classical model-theoretic results fail on descriptive frames.

The next lemma demonstrates that a subclass of finite frames can be defined with a single first-order sentence.[6]

Lemma 3.1. An infinite, irreflexive, linear order cannot be given the structure of a descriptive frame. Therefore, the subclass of finite, irreflexive, linear orders of the class of descriptive models can be defined by a single first-order sentence.

Proof. Suppose $\mathfrak{g} = (W, R, A)$ is a descriptive frame such that (W, R) is an infinite, irreflexive linear order. The $\subseteq$-decreasing sequence of sets $W \supseteq R[W] \supseteq \cdots \supseteq R^n[W] \supseteq \cdots$ must be non-empty for all n, otherwise $\mathfrak{g}$ would be a finite chain. As such, $C := \bigcap_n R^n[W]$ is closed and non-empty, because the space is topologically compact by assumption, the R-image of a closed set is closed [4,

[6] In [14, Lemma 3.6], we have attempted to present an intuitive, visual proof of this fact, but this proof requires a topological toolkit of "nets". Readers with a background in topology are encouraged to consult it but the proof included in this paper is purposefully elementary.

Proposition 5.83iii] and induction thus grants that $R^n[W]$ is closed for all n. By tightness and irreflexivity, for each $x \in C$ there must be a clopen set a_x such that $x \in a_x$ but $x \notin \langle R\rangle a_x$. As then

$$C \cap \bigcap_{x \in C} \langle R\rangle a_x = \emptyset,$$

topological compactness of $\mathfrak{g}$ gives a finite set $\{x_1, \ldots, x_k\}$ such that the finite intersection $C \cap \langle R\rangle a_{x_1} \cap \cdots \cap \langle R\rangle a_{x_k}$ is empty. Since the x_i are linearly ordered, there must be a least one, say x_1, so that $\langle R\rangle a_{x_1} \subseteq \langle R\rangle a_{x_i}$ for all i by transitivity, yielding $C \cap \langle R\rangle a_{x_1} = \emptyset$.

The fact that $\langle R\rangle a_{x_1}$ is clopen, hence compact, and the fact that $C = \bigcap_n R^n[W]$, imply that there is an $n \in \mathbb{N}$ such that $R^n[W] \cap \langle R\rangle a_{x_1} = \emptyset$, so that $x_1 \notin R^{n+1}[W]$, which contradicts $x_1 \in C$. Thus no such descriptive frame exists.

The irreflexive linear orders are definable in the single first-order sentence λ given by

$$\lambda := \qquad \underbrace{\forall x \forall y[(x \equiv y \vee Rxy) \leftrightarrow \neg Ryx]}_{\text{total, irreflexive, and antisymmetric}} \wedge \underbrace{\forall x \forall y \forall z[(Rxy \wedge Ryz) \rightarrow Rxz]}_{\text{transitive}}. \tag{1}$$

Since all finite irreflexive linear orders can be given the structure of a descriptive frame (with the powerset as a collection of admissible sets) the subclass of finite, irreflexive linear orders in the class of descriptive models can be defined by a single first-order sentence. $\square$

This lemma immediately implies the failure of the Compactness theorem for first-order logic.

Theorem 3.2. *The class of descriptive models is not compact over first-order logic.*

Proof. From Lemma 3.1, the Compactness theorem fails almost immediately. Taking λ to be Formula 1 and letting φ_n denote the existence of at least n distinct elements by

$$\varphi_n := \qquad \exists x_1 \cdots \exists x_n \bigwedge_{1 \le i < j \le n} \neg x_i \equiv x_j \qquad\qquad \text{where } n \in \mathbb{N},$$

Lemma 3.1 implies directly that $\{\lambda\} \cup \{\varphi_n\}_{n \in \mathbb{N}}$ is not satisfiable on the class of descriptive models. However, every finite subset is satisfied on a finite, irreflexive linear order, which can always be given the structure of a descriptive model with the powerset as collection of admissible sets. $\square$

With the failure of the Compactness theorem in mind, it is natural to wonder if some of its famous consequences fail, and indeed they do.

Theorem 3.3 (Failure of Beth definability theorem on descriptive models)**.** *Let R be a binary relation symbol and let P be a unary predicate symbol. Then there exists a formula φ that implicitly defines*[7] *P over the class $\mathscr{D}$ of descriptive models such that there is no formula $\psi \in \mathtt{FOL}(\{R, P\}, \emptyset)$ that explicitly defines P relative to φ*[8] *over $\mathscr{D}$.*

Proof. Let λ be Formula 1 defining the finite, irreflexive linear orders over the class of descriptive models as per Lemma 3.1. Then take φ to be the formula saying that R is an irreflexive linear order, the R-minimum does not have P and any two immediate successors disagree on P. Formally, φ is

$$\lambda \wedge \forall x \underbrace{[Px \leftrightarrow \exists y[Ryx \wedge \neg\exists z[Ryz \wedge Rzx] \wedge \neg Py]]}_{P \text{ iff an immediate predecessor lacks } P} \tag{2}$$

The models of λ are precisely finite, irreflexive linear orders that assign P to exactly the even points.

Suppose now that there were a formula $\psi(x) \in \mathtt{FOL}(\{R\}, \emptyset)$ that explicitly defined P. Then the formula ε stating that R is an irreflexive linear order and the maximum has the property ψ, given explicitly by

$$\lambda \wedge \forall x[\neg\exists y\ Rxy \to \psi(x)] \tag{3}$$

would be a formula in $\mathtt{FOL}(\{R\}, \emptyset)$ that characterises exactly the finite, even, irreflexive linear orders. However, sufficiently large linear orders are first-order equivalent as shown in [7, Example 2.3.6] with the Ehrenfeucht-Fraïssé method. It follows that no one formula can characterise the even linear orders and thus ψ cannot exist. $\square$

The Craig interpolation theorem fails on the class of descriptive models for essentially the same reason that the Beth definability theorem fails. Usually, the Beth definability theorem is stated as a consequence of the Craig interpolation theorem, so that the failure of former immediately implies failure of the latter, but we consider a direct proof to be informative.

Theorem 3.4 (Failure of Craig interpolation theorem on descriptive models)**.** *Let R be a binary relation symbol, let P_0, P_1 be unary predicate symbols, and let $\mathscr{D}_0, \mathscr{D}_1$, and $\mathscr{D}_{01}$ denote the class of descriptive models over predicate sets $\{P_0\}$, $\{P_1\}$, and $\{P_0, P_1\}$ respectively. Then there exist formulae $\varphi \in \mathtt{FOL}(\{R, P_0\}, \emptyset)$ and $\psi \in \mathtt{FOL}(\{R, P_1\}, \emptyset)$ such that $\varphi \models_{\mathscr{D}_{01}} \psi$, but there is no formula $\theta \in \mathtt{FOL}(\{R\}, \emptyset)$ such that $\varphi \models_{\mathscr{D}_0} \theta$ and $\theta \models_{\mathscr{D}_1} \psi$.*

Proof. Let φ state that R is an irreflexive linear order, P_0 occurs only on the even points as in Formula 2 in the proof of Theorem 3.3, as well as that the

[7] A formula φ in a language $\mathcal{L} \sqcup \{P\}$ *implicitly defines* a predicate P if every $\mathcal{L}$-structure $\mathfrak{A}$ has at most one extension to a $\mathcal{L} \sqcup \{P\}$-structure $\widetilde{\mathcal{A}}$ such that $\widetilde{\mathcal{A}} \models \varphi$.

[8] If φ in $\mathcal{L} \sqcup \{P\}$ implicitly defines P, then a formula $\psi(x)$ in $\mathcal{L}$ *explicitly defines* P relative to φ if $\varphi \models \forall x[\psi(x) \leftrightarrow Px]$.

maximum satisfies P_0 like Formula 3. Let ψ state that R is an irreflexive linear order and if P_1 occurs only on the even points, then the maximum has P_1.

The formula φ is true on the finite, even, irreflexive linear orders with P_0 on the even points. Similarly, for ψ to be true, a structure must be a finite, irreflexive linear order with P_1 true on an odd point or false on an even point, or it must be even. Since all structures satisfying φ are even, it follows that $\varphi \models_{\mathscr{D}_{01}} \psi$.

Suppose now for contradiction that there is an interpolant $\theta \in \mathtt{FOL}(\{R\}, \emptyset)$ such that $\varphi \models_{\mathscr{D}_0} \theta$ and $\theta \models_{\mathscr{D}_1} \psi$. Then in particular θ must be true on all finite, even, irreflexive linear R-orders. Again from [7, Example 2.3.6], this means that θ is true on all sufficiently large finite, irreflexive linear R-orders. But an odd linear order among these can be expanded with a predicate P_1 on the even points, from which it follows that $\theta \not\models_{\mathscr{D}_1} \psi$. This is a contradiction. $\square$

These two results use Lemma 3.1 to reduce a model-theoretic problem for the class of descriptive models to the corresponding result in finite model theory. This reasoning can resolve model-theoretic problems on descriptive models that have an analogue on finite models, but does not help when considering results that have no sensible corresponding statement in finite model theory.

For an example of such results, consider the celebrated Upward Löwenheim-Skolem theorem. It states that any first-order theory T with an infinite model has arbitrarily large models.

The Upward Löwenheim-Skolem theorem has no meaningful interpretation in finite model theory, because the hypothesis of the implication is vacuously false. However, as the theorem is an immediate consequence of the Compactness theorem, it is natural to wonder if it holds on the class of descriptive models on which this the compactness property fails. Indeed this class turns out to lack the Upward Löwenheim-Skolem property. In fact, it even fails when the theorem is restricted to speak only about formulae.

Theorem 3.5 (Failure of Upward Löwenheim-Skolem theorem for descriptive models)**.** *There exists a first-order formula that is satisfiable on an infinite descriptive model, but not satisfiable on any uncountable descriptive model.*

Proof. Let φ be a formula encoding a linear order with a reflexive minimum and no other reflexive points. That is, φ is the formula

$$\begin{aligned}
&\forall x \forall y [x \equiv y \vee (Rxy \leftrightarrow \neg Ryx)] && \text{(antisymmetric \& total)}\\
\wedge\ &\forall x \forall y \forall z [(Rxy \wedge Ryz) \to Rxz] && \text{(transitive)}\\
\wedge\ &\forall x \forall y [Rxx \wedge Ryy \to x \equiv y] && \text{(at most one reflexive point)}\\
\wedge\ &\exists x [Rxx \wedge \forall y\ Rxy] && \text{(there is a reflexive minimum).}
\end{aligned}$$

First, consider the Kripke frame $(\omega + 1, R)$ with Rab if and only if $a > b$ or $a = b = \omega$, equipped with the collection of admissible sets given by the finite subsets of $\mathbb{N}$ and cofinite subsets of $\mathbb{N}$ with ω. See Figure 1 for the underlying Kripke frame. This is easily checked to be a descriptive frame. Then any model based on this frame satisfies φ, because the relation is a linear order, has a reflexive minimum and no other reflexive points.

Now suppose that $\mathfrak{m} = (W, R, A, V)$ is a descriptive model of cardinality $|W| > \aleph_0$ whose underlying frame is a linear order with exactly one reflexive point v that is also its minimum. Because of transitivity and cardinality reasons, there is at least one point w with infinitely many successors that is not the minimum:

$$\aleph_0 < |W| = |R[v]| = |R[v] \setminus \{v\}| = \left| \bigcup_{x \neq v} \{x\} \cup R[x] \right| = \sup_{x \neq v} |\{x\} \cup R[x]| \,,$$

where the last equality follows from transitivity. So there is at least one $w \in W \setminus \{v\}$ such that $|\{w\} \cup R[w]| > \aleph_0$.

Now $R[w]$ is closed by [4, Proposition 5.83iv] and thus a compact subspace as a closed subset of a compact space, thereby inducing another infinite descriptive model that is a linear order. By Lemma 3.1, it follows that $R[w]$ contains another reflexive point, which means that $\mathfrak{m}$ contains at least two reflexive points, implying that $\mathfrak{m} \not\models \varphi$. $\square$

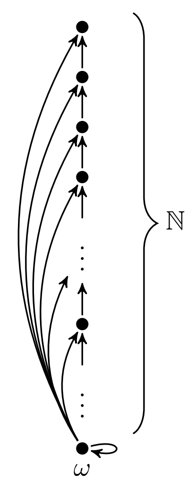

Figure 1: Schematic illustration of $(\omega + 1, R)$ for R the $\geq$-relation without reflexivity on the natural numbers. To avoid cluttering, not all arrows between the natural numbers are drawn, but the relation should be understood to be transitive.

4 The van Benthem characterisation theorem for descriptive models

In this section we prove the van Benthem characterisation theorem for the class of models over descriptive frames, stating that modal logic is the Kripke-bisimulation-invariant (or equivalently the Vietoris-bisimulation-invariant) fragment of first-order logic.

Theorem 4.1 (The van Benthem characterisation theorem for descriptive models)**.** *Let α be a first order formula in one free variable. Then the following are equivalent:*

1. *There exists a modal formula φ such that for all pointed, descriptive models $\mathfrak{m}, w$*

$$\mathfrak{m}, w \Vdash \varphi \textit{ if and only if } \mathfrak{m} \models \alpha[w].$$

2. *If $\mathfrak{m}, w$ and $\mathfrak{n}, v$ are two Kripke-bisimilar pointed, descriptive models, then*

$$\mathfrak{m} \models \alpha[w] \textit{ if and only if } \mathfrak{n} \models \alpha[v];$$

3. *If $\mathfrak{m}, w$ and $\mathfrak{n}, v$ are two Vietoris-bisimilar pointed, descriptive models, then*

$$\mathfrak{m} \models \alpha[w] \textit{ if and only if } \mathfrak{n} \models \alpha[v].$$

The strategy followed will be modelled after [21] and is visually represented in Figure 2. That is, the unravelling tree will be modified in Definitions 4.12

(Def 4.22) (Def 4.22)

$$\mathfrak{m}^{\otimes\kappa},(w,0) \quad \underline{\leftrightarrow} \quad \mathfrak{m},w \quad \underline{\leftrightarrow}_\ell \quad \mathfrak{n},v \quad \underline{\leftrightarrow} \quad \mathfrak{n}^{\otimes\kappa},(v,0)$$

(Def 4.2) (Def 4.2)

$$\mathfrak{T}(\mathfrak{m}^{\otimes\kappa}),((w,0)) \qquad\qquad \mathfrak{T}(\mathfrak{n}^{\otimes\kappa}),((v,0))$$

$$\widehat{\mathfrak{m}^{\otimes\kappa}},\widehat{w} \;\underline{\leftrightarrow}\; \widehat{\mathfrak{m}^{\otimes\kappa}} \uplus \widehat{\mathfrak{n}^{\otimes\kappa}},\widehat{w} \;\equiv_k^{\mathtt{FOL}}\; \widehat{\mathfrak{m}^{\otimes\kappa}} \uplus \widehat{\mathfrak{n}^{\otimes\kappa}},\widehat{v} \;\underline{\leftrightarrow}\; \widehat{\mathfrak{n}^{\otimes\kappa}},\widehat{v}$$

(Def 4.12) (Lem 4.26) (Def 4.12)

Figure 2: A visual representation of the argument that will be used to prove Theorem 4.1.

to remain in the class of descriptive models, while still Kripke-bisimilar (and Vietoris-bisimilar) to the original model. On these tree-like structures, an

Ehrenfeucht-Fraïssé-type argument (see [10] for more details on this technique) will show in Lemma 4.26 that bisimulation-invariance implies k-bisimulation-invariance for sufficiently large k. This will imply that any formula that is bisimulation-invariant on descriptive models is modally expressible. Assuming that α is bisimulation-invariant, $\mathfrak{m} \models \alpha[w]$, and $\mathfrak{m}, w \underline{\leftrightarrow}_\ell \mathfrak{n}, v$ for sufficiently large ℓ, the formula α can be followed clockwise around the diagram to conclude $\mathfrak{n} \models \alpha[v]$. The final conclusion of the proof can be found on page 23.

4.1 Unravelling for general frames

Towards proving the van Benthem theorem for descriptive models, it will be necessary to modify the notion of unravelling trees and unravelling forests to descriptive frames. To do this, first the unravelling forest of all finite paths needs to be given the structure of a general frame. The most canonical way of doing that, pulling back the admissible sets through the projection map π, does not offer a useful solution. The resulting frame would not inherit differentiatedness nor tightness. To accomplish this inheritance, $R^{\mathfrak{T}}$-closure and the collection $\mathcal{I}$ of paths of length 0 will be added.

Definition 4.2. Let $\mathfrak{g} = (\mathfrak{F}, A)$ be a general frame with underlying frame $\mathfrak{F} = (W, R)$. Define the *unravelling cover* $\mathfrak{T}(\mathfrak{g})$ of $\mathfrak{g}$ to be $\mathfrak{T}(\mathfrak{g}) := (\mathfrak{T}(\mathfrak{F}), A^{\mathfrak{T}})$, where $\mathfrak{T}(\mathfrak{F}) = (W^{\mathfrak{T}}, R^{\mathfrak{T}})$ defined by

$$\begin{aligned} W^{\mathfrak{T}} &:= \{(w_i)_{i \le n} \in W^{n+1} \mid n \in \mathbb{N}, \forall i < n : w_i R w_{i+1}\}; \\ R^{\mathfrak{T}} &:= \{((w_i)_{i \le n}, (w_i)_{i \le n+1}) : (w_i)_{i \le n+1} \in W^{\mathfrak{T}}\} \end{aligned}$$

is the forest of all finite paths in $\mathfrak{F}$ and with collection of admissible sets $A^{\mathfrak{T}}$, a subalgebra of $(\mathscr{P}(W^{\mathfrak{T}}), \langle R^{\mathfrak{T}} \rangle)$ defined through the surjective bounded morphism $\pi : \mathfrak{T}(\mathfrak{F}) \twoheadrightarrow \mathfrak{F}$ by the following recursive schema:

$$\begin{array}{lcll} & & \mathcal{I} := W^{\mathfrak{T}} \cap W \in A^{\mathfrak{T}}; & (1) \\ a \in A & \Longrightarrow & \pi^{-1}[a] \in A^{\mathfrak{T}}; & (2) \\ b \in A^{\mathfrak{T}} & \Longrightarrow & R^{\mathfrak{T}}[b] \in A^{\mathfrak{T}}; & (3) \\ b, b' \in A^{\mathfrak{T}} & \Longrightarrow & b \cup b', b \cap b' \in A^{\mathfrak{T}}. & (4) \end{array}$$

In a similar vein, for a general model $\mathfrak{m} = (\mathfrak{g}, V)$, the unravelling cover will be $\mathfrak{T}(\mathfrak{m}) := (\mathfrak{T}(\mathfrak{g}), \mathscr{P}^{\mathrm{op}}\pi \circ V)$, where $\mathscr{P}^{\mathrm{op}}\pi(a) = \pi^{-1}[a]$. For any point $w \in W$, the connected component of paths starting at w will be written as $\mathfrak{T}_w(\mathfrak{g})$ or $\mathfrak{T}_w(\mathfrak{m})$. ◀

Despite these extra additions, the resulting frame is always a general frame.

Proposition 4.3. Let $\mathfrak{g} = (W, R, A)$ be a general frame. Then $\mathfrak{T}(\mathfrak{g})$ is a general frame.

Proof. All to be checked is that $A^{\mathfrak{T}}$ is a field of sets and closed under the $\langle R^{\mathfrak{T}} \rangle$-operation. This is an elementary, but tedious induction. Closure under binary union and intersection is immediate. Closure under complements can by checked by induction on the construction of $A^{\mathfrak{T}}$. The complement of the initial points $\mathcal{I}^c = R^{\mathfrak{T}}[W^{\mathfrak{T}}]$ is the set of points with a predecessor, and is in $A^{\mathfrak{T}}$ because $W^{\mathfrak{T}} = \pi^{-1}[W]$ is in $A^{\mathfrak{T}}$, from which rule (3) gives $\mathcal{I}^c = R^{\mathfrak{T}}[W^{\mathfrak{T}}] \in A^{\mathfrak{T}}$.

Moreover, $(\pi^{-1}[a])^c = \pi^{-1}[a^c] \in A^{\mathfrak{T}}$ and if $b_0, b_1 \in A^{\mathfrak{T}}$ have $b_0^c, b_1^c \in A^{\mathfrak{T}}$, then $(b_0 \cup b_1)^c = b_0^c \cap b_1^c \in A^{\mathfrak{T}}$ and $(b_0 \cap b_1)^c = b_0^c \cup b_1^c \in A^{\mathfrak{T}}$.

The final case, the $R^{\mathfrak{T}}$-image of an admissible set, requires a minor observation about $R^{\mathfrak{T}}$: distinct points have disjoint image-sets. After all, two distinct paths cannot have the same extension. An alternative way of saying this is that each point has at most one $R^{\mathfrak{T}}$-predecessor. As such, we have for all $a \subseteq W^{\mathfrak{T}}$ that

$$(R^{\mathfrak{T}}[a])^c = R^{\mathfrak{T}}[a^c] \cup \mathcal{I} \in A^{\mathfrak{T}},$$

because each point with no predecessor in a either has a its unique predecessor in a^c or has no predecessor.

For closure under $\langle R^{\mathfrak{T}} \rangle$, note that every element in $A^{\mathfrak{T}}$ can be written as a finite union of finite intersections of elements of the form $\pi^{-1}[a]$ for $a \in A$ admissible, $R^{\mathfrak{T}}[b]$ for $b \in A^{\mathfrak{T}}$ or $\mathcal{I}$. This is evident for elements of the forms (1), (2) or (3) and will follow by induction on the construction for elements of the form (4).

As $\langle R^{\mathfrak{T}} \rangle$ distributes over unions, it is sufficient to show $\langle R^{\mathfrak{T}} \rangle$-closure for finite intersections. By induction on n, it will be shown that for $b_1, \ldots, b_n$ of the forms (1),(2) and (3), the set $\langle R^{\mathfrak{T}} \rangle (b_1 \cap \cdots \cap b_n)$ is in $A^{\mathfrak{T}}$.

For $n = 1$, note that $\langle R^{\mathfrak{T}} \rangle \mathcal{I} = \emptyset = \pi^{-1}[\emptyset] \in A^{\mathfrak{T}}$ and if $a \in A$ then $\langle R^{\mathfrak{T}} \rangle \pi^{-1}[a] = \pi^{-1}[\langle R \rangle a] \in A^{\mathfrak{T}}$ because π is a bounded morphism.

Moreover, $\langle R^{\mathfrak{T}} \rangle (R^{\mathfrak{T}}[b])$, again because each point has at most one predecessor, is the subset of b given by points with at least one successor. After all, $x \in \langle R^{\mathfrak{T}} \rangle (R^{\mathfrak{T}}[b])$ if and only if there is a y such that $y \in R^{\mathfrak{T}}[b]$ and $x R^{\mathfrak{T}} y$. That is equivalent y being a successor to x and having a predecessor in b, and since predecessors in this frame are unique, this predecessor must be x. So $\langle R^{\mathfrak{T}} \rangle (R^{\mathfrak{T}}[b]) = \langle R^{\mathfrak{T}} \rangle W^{\mathfrak{T}} \cap b = \pi^{-1}[\langle R \rangle W] \cap b \in A^{\mathfrak{T}}$.

Now suppose that $\langle R^{\mathfrak{T}} \rangle (b_1 \cap \cdots \cap b_n)$ when all b_i are of the forms (1),(2) or (3). Consider then $b_0 \cap b_1 \cap \cdots \cap b_n$ for $n \geq 1$. If one of the b_i is of the form $\mathcal{I}$, then it follows immediately that $\langle R^{\mathfrak{T}} \rangle (b_0 \cap b_1 \cap \cdots \cap b_n) \subseteq \langle R^{\mathfrak{T}} \rangle \mathcal{I} = \emptyset$. If one of them, without loss of generality b_0, is of the form $R^{\mathfrak{T}}[b']$, then observe that $x \in \langle R^{\mathfrak{T}} \rangle (R^{\mathfrak{T}}[b'] \cap b_1 \cap \cdots \cap b_n)$ if and only if $\exists y \in R^{\mathfrak{T}}[x] : y \in R^{\mathfrak{T}}[b'] \cap b_1 \cap \cdots \cap b_n$. By the uniqueness of predecessors, this is equivalent to having $x \in b'$ and $x \in \langle R^{\mathfrak{T}} \rangle (b_1 \cap \cdots \cap b_n)$. As the latter was admissible by the induction hypothesis, it follows that $\langle R^{\mathfrak{T}} \rangle (R^{\mathfrak{T}}[b'] \cap b_1 \cap \cdots \cap b_n) = b' \cap \langle R^{\mathfrak{T}} \rangle (b_1 \cap \cdots \cap b_n)$, which is admissible by closure under intersection.

Finally, if none of the b_i are of the form $\mathcal{I}$ or $R^{\mathfrak{T}}[b']$ for some $b' \in A^{\mathfrak{T}}$, then they must all be of the form $b_i = \pi^{-1}[a_i]$, from which it follows that

$$\begin{aligned}\langle R^{\mathfrak{T}}\rangle(b_0 \cap \dots \cap b_n) &= \langle R^{\mathfrak{T}}\rangle(\pi^{-1}[a_0] \cap \dots \cap \pi^{-1}[a_n]) = \langle R^{\mathfrak{T}}\rangle\pi^{-1}[a_0 \cap \dots \cap a_n] \\ &= \pi^{-1}[\langle R\rangle(a_0 \cap \dots \cap a_n)] \in A^{\mathfrak{T}},\end{aligned}$$

because $\langle R\rangle(a_0 \cap \dots \cap a_n) \in A$. $\square$

Now towards the main result of the paper, it is useful to find out which properties of the general frame are transferred to its unravelling cover.

Remark 4.4. For all $a \in A^{\mathfrak{T}}$ and $n \in \mathbb{N}$, the set $(R^{\mathfrak{T}})^n[a]$ is admissible. ◀

In order to obtain the van Benthem result for descriptive models, we will show that properties defining descriptive frames are preserved under this construction.

Proposition 4.5. If $\mathfrak{g} = (\mathfrak{F}, A)$ is differentiated, then so is $\mathfrak{T}(\mathfrak{g})$.

Proof. Let $\vec{x} = (x_i)_{i \le n}, \vec{y} = (y_j)_{j \le m} \in W^{\mathfrak{T}}$ be distinct paths. If $n > m$, then $\vec{y} \notin (R^{\mathfrak{T}})^n[W^{\mathfrak{T}}] \ni \vec{x}$. So assume $n = m$. Then their distinction must mean there is some $k \le n$ such that $x_k \ne y_k$. Then by differentiatedness of $\mathfrak{g}$ there is an $a \in A$ such that $y_k \notin a \ni x_k$. But then $(y_i)_{i \le k} \notin \pi^{-1}[a] \ni (x_i)_{i \le k}$, meaning $\vec{y} \notin (R^{\mathfrak{T}})^{n-k}[\pi^{-1}[a]] \ni \vec{x}$. $\square$

Proposition 4.6. If $\mathfrak{g} = (\mathfrak{F}, A)$ is differentiated, then $\mathfrak{T}(\mathfrak{g})$ is tight.

Proof. Let $(\vec{x}, \vec{y}) \notin R^{\mathfrak{T}}$. If the length $l(\vec{y})$ of $\vec{y}$ is not $l(\vec{x})+1$, then $\vec{y} \in (R^{\mathfrak{T}})^{l(\vec{y})}[\mathcal{I}]$, but $\vec{x} \notin \langle R^{\mathfrak{T}}\rangle(R^{\mathfrak{T}})^{l(\vec{y})}[\mathcal{I}] \subseteq (R^{\mathfrak{T}})^{l(\vec{y})-1}[\mathcal{I}]$, because all paths in $(R^{\mathfrak{T}})^n[\mathcal{I}]$ have length n.

If $n := l(\vec{y}) = l(\vec{x}) + 1$, then there must be a $k \le l(\vec{x})$ such that $x_k \ne y_k$. By assumption, $\mathfrak{g}$ is differentiated, so there exists an $a \in A$ with $x_k \notin a \ni y_k$. This implies that $\vec{y} \in (R^{\mathfrak{T}})^{n-k}[\pi^{-1}[a]] \in A^{\mathfrak{T}}$ but $\vec{x} \notin (R^{\mathfrak{T}})^{n-k-1}[\pi^{-1}[a]] \supseteq \langle R^{\mathfrak{T}}\rangle(R^{\mathfrak{T}})^{n-k}[\pi^{-1}[a]]$. $\square$

Although surprising, it may be understandable that tightness is not required, as the structure of the forest itself already separates unconnected points automatically.

Proposition 4.7. Let $\mathfrak{g} = (W, R, A)$ be a general frame and $\mathfrak{T}(\mathfrak{g})$ its unravelling cover. Take $\vec{x} \in W^{\mathfrak{T}}$. Then $\pi \restriction R^{\mathfrak{T}}[\vec{x}] : R^{\mathfrak{T}}[\vec{x}] \to R[\pi(\vec{x})]$ is a homeomorphism between the subspace topologies.

Proof. It is obviously a continuous bijection as $\pi^{-1}[a] \in A^{\mathfrak{T}}$ for all $a \in A$, which is the basis of the topology on $\mathfrak{g}$. To prove continuity of the inverse, it is sufficient to show that all basis elements in the subspace topology of $R^{\mathfrak{T}}[\vec{x}]$ are of the form $\pi^{-1}[a] \cap R^{\mathfrak{T}}[\vec{x}]$.

By induction on the recursion schema for $A^{\mathfrak{T}}$. It is obvious for the restriction of $\pi^{-1}[a]$ and $\mathcal{I}$. If $b \in A^{\mathfrak{T}}$, then

$$R^{\mathfrak{T}}[b] \cap R^{\mathfrak{T}}[\vec{x}] = \begin{cases} \emptyset & \text{if } \vec{x} \notin b, \\ R^{\mathfrak{T}}[\vec{x}] & \text{if } \vec{x} \in b, \end{cases}$$

by using again that distinct points have disjoint $R^{\mathfrak{T}}$-successor sets.

Finally, π^{-1} distributes over intersection and union, completing the induction. $\square$

Corollary 4.8. A general frame $\mathfrak{g}$ is image-compact if and only if $\mathfrak{T}(\mathfrak{g})$ is image-compact.

Proof. The previous proposition gives that the image-sets are homeomorphic, granting the corollary immediately. $\square$

Corollary 4.9. Let $\mathfrak{g}$ be a descriptive frame. Then $\mathfrak{T}(\mathfrak{g})$ is a differentiated, tight, and image-compact.

Proof. Differentiatedness and tightness were stated in Propositions 4.5 and 4.6. Note that descriptive frames are image-compact, because points are closed in Hausdorff spaces, the image of a closed set is closed in a descriptive frame [4, Proposition 5.83iii], and a closed subset of a compact space is also compact. From there, Corollary 4.8 finishes the proof. $\square$

4.2 The descriptive unravelling

The unravelling cover of a compact frame is not necessarily compact, as the following example demonstrates.

Example 4.10. Consider the descriptive frame consisting of a single reflexive point with the only field of sets possible. Its unravelling cover is $\mathbb{N}$ with the finite and cofinite sets as admissible sets, which is not compact.

This shows that while differentiatedness of $\mathfrak{g}$ implies $\mathfrak{T}(\mathfrak{g})$ is differentiated and tight, compactness may not be preserved. In fact, no collection of admissible sets can be constructed with which an unravelling forest of a descriptive frame with arbitrarily long paths is descriptive.

Proposition 4.11. Let $\mathfrak{g} = (\mathfrak{F}, A)$ be a descriptive frame. If the path lengths in $\mathfrak{g}$ are unbounded, then $\mathfrak{T}(\mathfrak{F})$ cannot be made into a descriptive frame.

Proof. For contraposition, let $(\mathfrak{T}(\mathfrak{F}), \mathcal{A})$ be descriptive. As reasoned before in the proof of Lemma 3.1, from [4, Proposition 5.83iii] and induction it follows that $(R^{\mathfrak{T}})^n[W^{\mathfrak{T}}]$ is closed for all n. Clearly, $\bigcap_{n\in\mathbb{N}}(R^{\mathfrak{T}})^n[W^{\mathfrak{T}}] = \emptyset$, as any one path is of finite length. By compactness, then, there exists an m such that $(R^{\mathfrak{T}})^m[W^{\mathfrak{T}}] = \emptyset$. Therefore, there exist no path of length m or more in $\mathfrak{g}$. $\square$

Like for the proof on finite models from [21], the unravelling must be modified to become descriptive. In principle, this could be done in the same way as was done in [21]. Putting the original frame at a sufficiently long distance from $\mathcal{I}$ would likely suffice for a reproduction of the argument. However, for descriptive frames, there exists an alternative construction that will be used in this paper: the descriptive completion using Jónsson-Tarski duality discussed in Definition 2.6.[9]

Definition 4.12. Let $\mathfrak{g}$ be a general frame. Then let its *descriptive unravelling* be the descriptive completion (Definition 2.6) of the unravelling cover (Definition 4.2) of $\mathfrak{g}$. Write

$$\widehat{\mathfrak{g}} := ((\mathfrak{T}(\mathfrak{g}))^*)_*$$

to abbreviate. ◀

This will turn out to be a very well-behaved construction for descriptive frames and will be key to the theorem.

Lemma 4.13. Let $\mathfrak{g} = (W, R, A)$ be image-compact. Let $F_x, F \in \mathrm{Uf}\ A$, where F_x is the ultrafilter generated by x. Then in the descriptive completion $F_x R_* F$ if and only if $F = F_y$ for some $y \in R[x]$.

Proof. For the implication from right to left, assume that $y \in R[x]$. To prove that $F_x R_* F_y$, let $a \in F_y$. Then $y \in a$. From xRy it follows that $x \in \langle R \rangle a$, yielding $\langle R \rangle a \in F_x$. Since a was arbitrary, this holds for all $a \in F_y$, so that $F_x R_* F_y$.

For the implication from left to right, assume $F_x R_* F$. By definition, if $a \in F$ then $\langle R \rangle a \in F_x$, implying $x \in \langle R \rangle a$. Therefore, there exists an $x' \in a$ such that xRx'.

So for every $a \in F$, we have $a \cap R[x] \neq \emptyset$. Since F is closed under finite intersections, we find that $\{a \cap R[x] : a \in F\}$ has the finite intersection property. By compactness of $R[x]$, the set $\bigcap\{a \cap R[x] : a \in F\} = R[x] \cap \bigcap F$ is non-empty. Therefore, there exists a $y \in R[x]$ such that for all $a \in F$ we have $y \in a$. So $F = F_y$ for some $y \in R[x]$, because it is an ultrafilter. □

Remark 4.14. One might have expected tightness to show up in the proposition above to prove that $F_x R_* F_y$ implies xRy, but tightness is an immediate consequence from image-compactness and differentiability, so it may be reasonable to expect that image-compactness is strong enough to prove something not quite as strong as tightness. ◀

Proposition 4.15. Let $\mathfrak{g}$ be a differentiated and image-compact frame. Then $\mathfrak{g} \rightarrowtail (\mathfrak{g}^*)_*$ is a generated subframe[10] through a topological embedding $\iota_{\mathfrak{g}} : x \mapsto F_x$. That is, $\mathfrak{g}$ is homeomorphic to its image under $\iota_{\mathfrak{g}}$.

Proof. Lemma 4.13 gives immediately that it is a bounded morphism. From the fact that $\mathfrak{g}$ is differentiated, it follows that $\iota_{\mathfrak{g}}$ injective, making $\mathfrak{g}$ a generated

[9] Actually, the construction is quite natural in that it admits several potential equivalent definitions, including an explicit construction and a topological definition [14, Definition 2.86, Definition 5.14] as well as a definition through a universal property, but these will not be needed in this paper.

[10] Generated subframes are given by injective bounded morphisms

subframe. To see that it is a homeomorphism on its image, let a be an admissible set on $\mathfrak{g}$. Then

$$\iota_{\mathfrak{g}}(x) = F_x \in \widehat{a} \iff a \in F_x \iff x \in a,$$

so that $\iota_{\mathfrak{g}}$ and $\iota_{\mathfrak{g}}^{-1}$ preserve the basis elements of the topology, ensuring continuity for both it and its inverse restricted to the image. $\square$

Theorem 4.16. *Let $\mathfrak{g}$ be a descriptive frame. Then $\widehat{\iota} : \mathfrak{T}(\mathfrak{g}) \rightarrowtail \widehat{\mathfrak{g}}$ continuously.*

Proof. Note that descriptive frames are in particular, image-compact and differentiated, so Corollary 4.8 and Proposition 4.5 gives that $\mathfrak{T}(\mathfrak{g})$ is image-compact and differentiated. Proposition 4.15 then gives the theorem. $\square$

In fact, an even stronger claim is true.

Theorem 4.17. *Let $\mathfrak{g}$ be a descriptive frame. Then $\widehat{\mathfrak{g}}_\# = \mathfrak{T}(\mathfrak{g})_\# \uplus \mathfrak{L}$ for some unspecified frame of limit points $\mathfrak{L}$,[11] where the $\#$-operation takes the underlying Kripke frame of a general frame.*

Proof. Let $\mathfrak{g} = (W, R, A)$. From Theorem 4.16, it is sufficient to show that two points in $\widehat{\mathfrak{g}}$ can only be $(R^{\mathfrak{T}})_*$-related if they are both inside or both outside $\mathfrak{T}(\mathfrak{g})_\#$. Theorem 4.16 implies that if w is in the image of $\widehat{\iota} : \mathfrak{T}(\mathfrak{g}) \rightarrowtail \widehat{\mathfrak{g}}$, then the $(R^{\mathfrak{T}})_*$-successor set of w is, too. To complete the theorem, the predecessor set has to be, as well. This means that if $F(R^{\mathfrak{T}})_* F_x$ for some ultrafilter F and the ultrafilter F_x generated by x, then $F = F_y$ for $y \in (R^{\mathfrak{T}})^{-1}[x]$.

Towards contraposition, assume that $F \neq F_y$ for any $y \in (R^{\mathfrak{T}})^{-1}[x]$. In $\mathfrak{T}(\mathfrak{g})$, each point has at most one predecessor, so $(R^{\mathfrak{T}})^{-1}[x] \subseteq \{y\}$ for some y. In particular, this means there exists some $a \in F \subseteq A^{\mathfrak{T}}$ such that $x \notin R^{\mathfrak{T}}[a]$, either because it has no predecessor or because $y \notin a$. By construction of $A^{\mathfrak{T}}$, the set $R^{\mathfrak{T}}[a]$ is in $A^{\mathfrak{T}}$, so that $a \subseteq [R^{\mathfrak{T}}]R^{\mathfrak{T}}[a] \in F$ by monotonicity of filters. Since $x \notin R^{\mathfrak{T}}[a]$, also $R^{\mathfrak{T}}[a] \notin F_x$, implying that $(F, F_x) \notin (R^{\mathfrak{T}})_*$. $\square$

As such, there is is an isomorphic copy of $\mathfrak{T}(\mathfrak{g})_\#$ in $\widehat{\mathfrak{g}}_\#$. The next step is to upgrade the descriptive unravelling to a descriptive model.

Corollary 4.18. Let $\mathfrak{m} = (\mathfrak{g}, V)$ be a descriptive model. Define $\widehat{\mathfrak{m}} := (\widehat{\mathfrak{g}}, \widehat{V})$ with $\widehat{V}(p) = \widehat{\pi^{-1}[V(p)]}$. Then the relation $\pi^T ; \widehat{\iota}$ is a Kripke bisimulation between $\mathfrak{m}$ and $\widehat{m}$, where the functions π and $\widehat{\iota}$ are to be viewed as binary relations, ; denotes composition of the relations and $-^T$ takes the inverse of a relation. Therefore its closure is a Vietoris bisimulation.

Proof. Both π and $\widehat{\iota}$ are bounded morphisms, so they satisfy the back and forth conditions by construction. The propositional requirements is satisfied because

$$\begin{aligned} \pi(\vec{x}) \in V(p) &\iff \vec{x} \in \pi^{-1}[V(p)] \iff \pi^{-1}[V(p)] \in F_{\vec{x}} \\ &\iff F_{\vec{x}} \in \widehat{\pi^{-1}[V(p)]} = \widehat{V}(p). \end{aligned}$$

The final remark is then given by [2, Theorem 5.2]. $\square$

[11] See [14, Chapter 5] for the structure of these limit points.

4.3 Preservation under finite bisimulations

The previous section provides a tool with which to show invariance under bisimulation implies invariance under some finite bisimulation. This tool will now be used to achieve this through Ehrenfeucht-Fraïssé methods. To this end, there is a final combinatorial construction that will prove useful: a duplication process. It will be useful to copy points in a manner that preserves compactness.

Definition 4.19. Let A and B be fields of sets over universes X and Y. Write $A \otimes B$ for the field of sets over the universe $X \times Y$ generated by $\{a \times b \mid a \in A, b \in B\}$. ◀

Proposition 4.20. Let A and B be fields of sets generating topological spaces $\mathcal{X}$ and $\mathcal{Y}$ as bases. Then $A \otimes B$ is a basis for the product space $\mathcal{X} \times \mathcal{Y}$.

Proof. To see that $A \otimes B$ generates topology that is at least as fine, note that the basis of $\mathcal{X} \times \mathcal{Y}$ consists of all $U \times V$ where U and V are open subsets of $\mathcal{X}$ and $\mathcal{Y}$ respectively. This means that $U = \bigcup_{\eta \in I} a_\eta$ and $V = \bigcup_{\xi \in J} b_\xi$ for $a_\eta \in A$ and $b_\xi \in B$ and index sets I and J. But then it is immediate that

$$U \times V = \left(\bigcup_{\eta \in I} a_\eta \right) \times \left(\bigcup_{\xi \in J} b_\xi \right) = \bigcup_{\eta \in I, \xi \in J} a_\eta \times b_\xi$$

is generated by $A \otimes B$.

To see that it is at least as coarse, remark that all elements of $A \otimes B$ are products of sets open in $\mathcal{X}$ with open sets of $\mathcal{Y}$. □

Corollary 4.21. Let A and B be fields of sets.

1. If A and B are compact, then $A \otimes B$ is compact;
2. If A and B are differentiated, then $A \otimes B$ is differentiated.

Proof. This follows immediately from Proposition 4.20 and the fact that the product of two compact spaces is compact and the product of totally separated spaces is totally separated. □

Definition 4.22. Let $\mathfrak{g} = (W, R, A)$ be a general frame and let $\mathcal{F}$ be a field of sets over a universe X. Define the *$\mathcal{F}$-multiplier* of $\mathfrak{g}$ by

$$\begin{aligned} R \otimes X &:= \{((w, x_1), (v, x_2)) \in (W \times X)^2 \mid (w, v) \in R, x_1, x_2 \in X\}; \\ \mathfrak{g}^{\otimes \mathcal{F}} &:= (W \times X, R \otimes X, A \otimes \mathcal{F}). \end{aligned}$$

If $\mathfrak{m} = (\mathfrak{g}, V)$ is a general model, then define $V(\cdot) \times X$ by $p \mapsto V(\cdot) \times X)$ and

$$\mathfrak{m}^{\otimes \mathcal{F}} := (\mathfrak{g}^{\otimes \mathcal{F}}, V(\cdot) \times X),$$

which will be called the $\mathcal{F}$-multiplier of $\mathfrak{m}$. ◀

Remark 4.23. There is an obvious surjective continuous bounded morphism $\pi_0 : \mathfrak{F}^{\otimes X} \twoheadrightarrow \mathfrak{F}$ given by projection on the first coordinate. ◀

Lemma 4.24. Let $\mathfrak{g} = (W, R, A)$ be a descriptive frame and let $\mathcal{F}$ over X be a compact and differentiated field of sets. Then $\mathfrak{g}^{\otimes \mathcal{F}}$ is a descriptive frame.

Proof. Compactness and differentiatedness follow immediately from Corollary 4.21. To see tightness, let $((w, x_1), (v, x_2)) \notin R \otimes X$. Then $(w, v) \notin R$. From tightness of $\mathfrak{g}$ follows the existence of $a \in A$ such that $v \in a$ but $w \notin \langle R \rangle a$. Then in particular, $(v, x_2) \in a \times X$, but $(w, x_1) \notin (\langle R \rangle a) \times X$. It is a general frame in the first place, because

$$\begin{aligned}
\langle R \otimes X \rangle (a \times X) &= \{\vec{s} \in W \times X \mid \exists \vec{t} \in a \times X : \vec{s}(R \otimes X)\vec{t}\} \\
&= \{(s, x) \in W \times X \mid \exists (t, y) \in a \times X : sRt\} \\
&= \{s \in W \times X \mid \exists t \in a : sRt\} \times X = (\langle R \rangle a) \times X,
\end{aligned}$$

completing the proof. □

As mentioned before, this construction will be used to apply Ehrenfeucht-Fraïssé method. Multiple constructions are conceivable, but the approach taken here is adopted for its convenience. It will be inspired by Hanf's Lemma [11, Lemma 2.3]. Like Hanf's Lemma, it relies on the notion of the Gaifman neighbourhood.

Definition 4.25. Let $\mathfrak{F} = (W, R)$ be a frame, $\mathfrak{M} = (\mathfrak{F}, V)$ a model and $S \subseteq W$. Then the *Gaifman neighbourhood of size ℓ of S* is defined recursively by

$$\begin{aligned}
N_0^{\mathfrak{F}}(S) &= S; \\
N_{\ell+1}^{\mathfrak{F}}(S) &= N_\ell^{\mathfrak{F}}(S) \cup \langle R \rangle N_\ell^{\mathfrak{F}}(S) \cup R\big[N_\ell^{\mathfrak{F}}(S)\big]; \\
\mathcal{N}_\ell^{\mathfrak{F}}(S) &= (N_\ell^{\mathfrak{F}}(S), R \restriction N_\ell^{\mathfrak{F}}(S)) && \text{where } R \restriction A := R \cap A \times A; \\
\mathcal{N}_\ell^{\mathfrak{M}}(S) &= (\mathcal{N}_\ell^{\mathfrak{F}}(S), V \cap N_\ell^{\mathfrak{F}}(S)),
\end{aligned}$$

where $V \cap N_\ell^{\mathfrak{F}}(S) := \Big(p \mapsto V(p) \cap N_\ell^{\mathfrak{F}}(S)\Big)$. That is, the ℓ-neighbourhood of S is the set of points that can be reached in ℓ steps along or against R. The letter N is used for the set, and $\mathcal{N}$ is used for the subframe and the submodel. ◀

That is, the ℓ-neighbourhood of a set S is the collection of sets that can be reached in at most ℓ steps forwards or backwards along R from S.

Lemma 4.26. Let $\mathfrak{M} = X \uplus \biguplus_{\eta \in I} B_\eta$ and $\mathfrak{N} = X \uplus \biguplus_{\rho \in J} C_\rho$ be two (general) models such that

- $B_\eta \cong B_\xi$ for all $\eta, \xi \in I$;
- $C_\rho \cong C_\pi$ for all $\rho, \pi \in J$;
- I and J are infinite.

Taking $\ell = 3^n$, suppose that for each point w in B_η there exists a point v in a C_ρ such that $\mathcal{N}_\ell^{\mathfrak{M}}(w) \cong \mathcal{N}_\ell^{\mathfrak{N}}(v)$ and vice versa. Then for w_0 in $\mathfrak{M}$ and v_0 in $\mathfrak{N}$

$$\mathcal{N}_\ell^{\mathfrak{M}}(w_0) \cong \mathcal{N}_\ell^{\mathfrak{N}}(v_0) \implies \mathfrak{M}, w_0 \equiv_n \mathfrak{N}, v_0.$$

Proof. The proof uses the Ehrenfeucht-Fraïssé method. The reader is referred to [10] for an exposition of the method.

By induction on the number of rounds played so far, it will be shown that Duplicator can counter any move by Spoiler. To this end, write $\ell(k) := 3^{n-k}$.

More precisely, let the induction hypothesis denote that for any $(a_i)_{i<k}$ and $(b_i)_{i<k}$ such that

$$X \cup \mathcal{N}_{\ell(k)}\Big(\{w_0\} \cup \{a_i\}_{i<k}\Big) \cong X \cup \mathcal{N}_{\ell(k)}\Big(\{v_0\} \cup \{b_i\}_{i<k}\Big)$$
$$\text{where } \begin{cases} a_i \mapsto b_i, \\ w_0 \mapsto v_0, \\ x \mapsto x & \text{if } x \in X. \end{cases}$$

a move a_k can be countered with a move b_k such that the above condition holds with k replaced by $k+1$. In particular, there will be a local isomorphism between the elements chosen. From symmetry, the response of an a_k to a b_k can be given similarly. Inductively performing this until $k = n$ then gives victory for Duplicator.

Suppose that there have been k turns and the inductive hypothesis holds. When Spoiler makes a move a_k, there are two cases to consider:

1. $a_k \in N_{2\cdot\ell(k+1)}\Big(\{w_0\} \cup \{a_i\}_{i<k}\Big) \cup X$,
2. $a_k \notin N_{2\cdot\ell(k+1)}\Big(\{w_0\} \cup \{a_i\}_{i<k}\Big) \cup X$

In the former case, let $\theta : X \cup N_{\ell(k)}(\{a_i\}_{i<k}) \to X \cup N_{\ell(k)}(\{b_i\}_{i<k})$ be the isomorphism assumed in the induction hypothesis and let $b_k = \theta(a_k)$. Observe that $2 \cdot 3^{n-k-1} + 3^{n-k-1} = 3^{n-k}$, so $N_{\ell(k-1)}(\{a_i\}_{i\le k}) \subseteq N_{\ell(k)}(\{a_i\}_{i<k})$ and the same for b. It follows immediately that the restriction of θ is again an isomorphism between these two smaller neighbourhoods.

In the second case, remark that $N_{\ell(k)}(\{b_i\}_{i<k})$ can only intersect at most k connected components. Because J was infinite per assumption, there is a $\rho \in J$ such that $N_{\ell(k)}(\{b_i\}_{i<k}) \cap C_\rho = \emptyset$. By assumption, there is a π such that

$$\mathcal{N}_{\ell(0)}(a_k) \cong \mathcal{N}_{\ell(0)}(v) \subseteq C_\pi \cong C_\rho$$

for some v. Therefore, there is a b_k in C_ρ (namely the image of v under the isomorphism directly above) such that $\mathcal{N}_{\ell(0)}(a_k) \cong \mathcal{N}_{\ell(0)}(b_k)$. Choosing it, one obtains

$$\begin{aligned}
\mathcal{N}_{\ell(k+1)}(a_k) &\cong \mathcal{N}_{\ell(k+1)}(b_k);\\
X \cup \mathcal{N}_{\ell(k+1)}(\{a_i\}_{i<k}) &\cong X \cup \mathcal{N}_{\ell(k+1)}(\{a_i\}_{i<k}) \qquad \text{through restriction of } \theta;\\
X \cup \mathcal{N}_{\ell(k+1)}(\{a_i\}_{i\leq k}) &= \mathcal{N}_{\ell(k+1)}(a_k) \uplus \big(X \cup \mathcal{N}_{\ell(k+1)}(\{a_i\}_{i<k})\big)\\
&\cong \mathcal{N}_{\ell(k+1)}(b_k) \uplus \big(X \cup \mathcal{N}_{\ell(k+1)}(\{b_i\}_{i<k})\big)\\
&= X \cup \mathcal{N}_{\ell(k+1)}(\{b_i\}_{i\leq k}),
\end{aligned}$$

where $\uplus$ denotes disjoint union of models and the isomorphism on the different disjoint components is preserved, so that it is still the identity on X, and sends the a_i to the b_i.

Symmetry of the models $\mathfrak{M}$ and $\mathfrak{N}$ guarantees that the exact same argument provides a response a_k to a move b_k by Spoiler.

At the end, this gives Duplicator an isomorphism

$$X \cup \mathcal{N}_{\ell(n)}\Big(\{w_0\} \cup \{a_i\}_{i<n}\Big) \cong X \cup \mathcal{N}_{\ell(n)}\Big(\{w_0\} \cup \{a_i\}_{i<n}\Big)$$

which, using $\ell(n) = 1$, can be restricted to a local isomorphism

$$\{w_0\} \cup \{a_i\}_{i<n} \cong \{v_0\} \cup \{b_i\}_{i<n},$$

winning the game for Duplicator. □

This final lemma now permits a proof of the van Benthem Characterisation Theorem for the class of models on descriptive frames.

Proof of Theorem 4.1. Remark that the equivalence (2) $\iff$ (3) follows immediately from Theorem 2.11.

The implication (1) $\implies$ (2) is a well-known property of Kripke bisimulations and modal logic [4, Theorem 2.20]. Modal formulae are bisimulation-invariant, so any first-order formula equivalent to one, is also bisimulation-invariant.

The only interesting implication is the implication towards (1). Towards the implication (2) $\implies$ (1), it will turn out to be sufficient that Kripke-bisimulation-invariance implies finite-bisimulation-invariance.

Let $\alpha(x)$ be a bisimulation-invariant formula in one free variable with quantifier depth $\mathrm{q}(\alpha)$. Write $\ell = 2 \cdot 3^{\mathrm{q}(\alpha)}$ and assume that $\mathfrak{m}, w$ and $\mathfrak{n}, v$ are pointed descriptive models with respective universes W and $\widetilde{W}$ such that

- $\mathfrak{m} \models \alpha[w]$;
- $\mathfrak{m}, w \mathrel{\underline{\leftrightarrow}_\ell} \mathfrak{n}, v$ over the propositional variables corresponding to the predicates occurring in α.

Let κ be an infinite cardinal greater than those of the universes of $\mathfrak{m}$ and $\mathfrak{n}$. Then the order topology is compact and totally separated on the ordinal number

$\kappa+1$. Write K for the field of clopens of this topology. Remark 4.23 gives that $\mathfrak{m}^{\otimes K},(w,0) \underline{\leftrightarrow}_\ell \mathfrak{n}^{\otimes K},(v,0)$ and Corollary 4.18 implies

$$\widehat{\mathfrak{m}^{\otimes K}}, F_{((w,0))} \underline{\leftrightarrow}_\ell \widehat{\mathfrak{n}^{\otimes K}}, F_{((v,0))}.$$

Write $\widehat{w} := F_{((w,0))}$ and $\widehat{v} := F_{((v,0))}$ to avoid unnecessarily complicated expressions. Observe that the bisimulations provided by Remark 4.23 and Corollary 4.18, together with the bisimulation-invariance of α imply that $\widehat{m^{\otimes K}} \models \alpha[\widehat{w}]$. Showing now that $\widehat{\mathfrak{n}^{\otimes K}} \models \alpha[\widehat{v}]$ would, through the same bisimulations, then show $\mathfrak{n} \models \alpha[v]$, as desired.

Following Theorem 4.17, the models $\widehat{\mathfrak{m}^{\otimes K}}$ and $\widehat{\mathfrak{n}^{\otimes K}}$ can be written as

$$\begin{aligned}
\widehat{\mathfrak{m}^{\otimes K}} &= \mathfrak{L} \uplus \mathfrak{T}(\mathfrak{m}^{\otimes K}) \overset{*}{=} \mathfrak{L} \uplus \biguplus_{\substack{\eta\in\kappa+1\\ w\in W}} \mathfrak{T}_{(w,\eta)}(\mathfrak{m}^{\otimes K});\\
\widehat{\mathfrak{n}^{\otimes K}} &= \Lambda \uplus \mathfrak{T}(\mathfrak{n}^{\otimes K}) \overset{*}{=} \Lambda \uplus \biguplus_{\substack{\eta\in\kappa+1\\ v\in \widetilde{W}}} \mathfrak{T}_{(v,\eta)}(\mathfrak{n}^{\otimes K});\\
\widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}} &= (\mathfrak{L} \uplus \Lambda) \uplus \biguplus_{\eta\in\kappa+1} \Big(\biguplus_{w\in W} \mathfrak{T}_{(w,\eta)}(\mathfrak{m}^{\otimes K}) \uplus \biguplus_{v\in\widetilde{W}} \mathfrak{T}_{(v,\eta)}(\mathfrak{n}^{\otimes K}) \Big),
\end{aligned}$$

where the equalities marked by $*$ follow from the fact that two paths can only be related in the tree if one extends the other, and two paths with different starting points cannot extend one another. Recall from Definition 4.2 that $T_x(\mathfrak{M})$ is the submodel of paths starting at x. The submodels $\mathfrak{L}$ and Λ are unspecified frames of limit points.

For any $\eta, \xi \in \kappa+1$,

$$\biguplus_{w\in W} \mathfrak{T}_{(w,\eta)}(\mathfrak{m}^{\otimes K}) \uplus \biguplus_{v\in\widetilde{W}} \mathfrak{T}_{(v,\eta)}(\mathfrak{n}^{\otimes K}) \cong \biguplus_{w\in W} \mathfrak{T}_{(w,\xi)}(\mathfrak{m}^{\otimes K}) \uplus \biguplus_{v\in\widetilde{W}} \mathfrak{T}_{(v,\xi)}(\mathfrak{n}^{\otimes K})$$

through the map of switching the initial point. Moreover,

$$\widehat{\mathfrak{m}^{\otimes K}}, \widehat{w} \underline{\leftrightarrow} \widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}}, \widehat{w} \qquad \Longrightarrow \qquad \widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}} \models \alpha[\widehat{w}]$$

and

$$\widehat{\mathfrak{n}^{\otimes K}}, \widehat{v} \underline{\leftrightarrow} \widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}}, \widehat{v}.$$

Therefore, if $\mathcal{N}_\ell^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) \cong \mathcal{N}_\ell^{\widehat{\mathfrak{n}^{\otimes K}}}(\widehat{v})$, applying Lemma 4.26 for

$$\begin{aligned}
\mathfrak{M} = \mathfrak{N} &:= \widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}};\\
I = J &:= \kappa+1;\\
B_\eta = C_\eta &:= \biguplus_{w\in W} \mathfrak{T}_{(w,\eta)}(\mathfrak{m}^{\otimes K}) \uplus \biguplus_{v\in\widetilde{W}} \mathfrak{T}_{(v,\eta)}(\mathfrak{n}^{\otimes K});\\
X &= \mathfrak{L} \uplus \Lambda
\end{aligned}$$

will prove that

$$\mathfrak{M}, \widehat{w} \equiv_{\mathrm{q}(\alpha)} \mathfrak{N}, \widehat{v}.$$

This implies $\widehat{\mathfrak{m}^{\otimes K}} \uplus \widehat{\mathfrak{n}^{\otimes K}} \models \alpha[\widehat{v}]$, from which the bisimulation above provides $\widehat{\mathfrak{n}^{\otimes K}} \models \alpha[\widehat{v}]$, which in turn shows $\mathfrak{n} \models \alpha[v]$ from the previous bisimulations. This implies that α is preserved under finitary bisimulations. As bisimilarity up to depth ℓ over finitely many propositional variables is identical to modal equivalence up to depth ℓ ([21, Proposition 1]), this means that α is characterised by some collection of theories of formulae of finite modal depth. As there are finitely many such formulae, there is a modal formula of depth up to ℓ equivalent to α, concluding the proof.

This final claim of isomorphism of neighbourhoods will be achieved by showing inductively that there exists a sequence of isomorphisms $(f_i)_{i\leq\ell}$ with

$$\begin{aligned}
&f_i : \mathcal{N}_i^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) \cong \mathcal{N}_i^{\widehat{\mathfrak{n}^{\otimes K}}}(\widehat{v}); \\
&f_i \upharpoonright \mathcal{N}_j^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) = f_j && \text{for } j \leq i; \\
&N_i^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) \ni x \leftrightsquigarrow_{\ell-i} f_i(x) \in N_i^{\widehat{\mathfrak{n}^{\otimes K}}}(\widehat{v}) && \text{over the predicates in } \alpha.
\end{aligned}$$

Clearly, the map $f_0 : \mathcal{N}_0^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) \to \mathcal{N}_0^{\widehat{\mathfrak{n}^{\otimes K}}}(\widehat{v})$ is an isomorphism, since both consist of a single irreflexive point and must satisfy the same propositional variables, because $\mathfrak{m}^{\otimes K}, \widehat{w} \,\underline{\leftrightarrow}_\ell\, \mathfrak{n}^{\otimes K}, \widehat{v}$, so they satisfy the same propositional variables, implying that f_0 is an isomorphism. Moreover, through [21, Proposition 1], it implies that $\widehat{w} \leftrightsquigarrow_\ell \widehat{v}$.

Now suppose that $f_i : \mathcal{N}_i^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) \to \mathcal{N}_i^{\widehat{\mathfrak{n}^{\otimes K}}}(\widehat{v})$ is an isomorphism. Since $\widehat{w}$ is the root of its treelike connected component, if R is the relation on $\widehat{\mathfrak{m}^{\otimes K}}$, then

$$N_j^{\widehat{\mathfrak{m}^{\otimes K}}}(\widehat{w}) = \bigcup_{k\leq j} R^k[\widehat{w}],$$

so the isomorphism f_{i+1} only needs to extend f_i on the successors of $R^i[\widehat{w}]$. Because the connected component under consideration is a tree, no two points share successors. Therefore, the successors of each such point may be considered separately.

Let $x \in R^i[\widehat{w}]$. Each theory of modal depth $\ell - i - 1$ over the predicates in α has a single characterising formula, as there are finitely many such formulae. Write Σ for the set of these characterising formulae. For each $\varphi \in \Sigma$, if $\mathfrak{m}^{\otimes K}, x \Vdash \Diamond\varphi$, then this type occurs κ many times in $R[x]$ and otherwise it occurs 0 times, because if it occurs once, then it must occur in all the K-duplicates. The same reasoning applies to $f_i(x)$, so because they agree on each $\Diamond\varphi$ per assumption, for each theory x and $f_i(x)$ either both have κ many successors with that theory, or none.

It follows that for every $\varphi \in \Sigma$, there exists a bijection g^x_φ between the successors of x satisfying φ and the successors of $f_i(x)$ satisfying φ. Choosing one such g^x_φ for each φ allows the construction of

$$f_{i+1} = f_i \sqcup \bigsqcup_{\substack{x \in R^i[\widehat{w}] \\ \varphi \in \Sigma}} g^x_\varphi,$$

this preserves all new relations and predicates. Moreover, $y \in R[x]$ with $x \in R^i[\widehat{w}]$ has modal theory characterised by $\varphi \in \Sigma$, and therefore by construction $y \leftrightsquigarrow_{\ell-i-1} g^x_\varphi(y)$. Since this also held for f_i by induction hypothesis, the inductive condition is satisfied again. $\square$

5 Conclusions and future work

The main result of this paper is the validity of the van Benthem characterisation theorem on the class of models based on descriptive frames. We also showed the failures of the Compactness theorem, the Upward Löwenheim-Skolem theorem, the Beth definability theorem, and the Craig interpolation theorem on this class of models.

We will now briefly discuss potential directions for future work. As discussed in the introduction, many generalisations and adaptations exist of the classical van Benthem characterisation theorem. Deserving special mention is the Janin-Walukiewicz theorem [15]. This theorem is an analogue of the van Benthem characterisation theorem for the modal μ-calculus. It states that the modal μ-calculus is the bisimulation-invariant fragment of monadic first-order logic. The proof of this theorem also relies centrally on unravelling trees and the convenient properties of trees for game-theoretic semantics. In [3], a subclass of descriptive frames is designed to allow interpretation of the modal μ-calculus. We leave it as an open problem whether an analogue of the Janin-Walukiewicz theorem can be proved for these descriptive μ-frames. We only note that for this task one could try to work with an appropriate version of descriptive unravellings defined in this paper.

The coalgebraic generalisations of the van Benthem characterisation theorem for finite frames are inspired by the view of Kripke frames as coalgebras for the powerset functor on the category of sets. Similarly, the finite Kripke frames can be viewed as coalgebras for the powerset functor in the category of finite sets. This then leads to a strategy based on the pseudotrees as in [21] to obtain the van Benthem characterisation theorem for finite supported coalgebras. As was shown in this paper, descriptive frames are model-theoretically highly similar to finite frames. As descriptive frames are coalgebras for the Vietoris functor on the category of Stone spaces, one could consider combining the constructions from this paper with the approach used in [22]. There again, a type of pseudotree is introduced to apply arguments from finite model theory to obtain the result. Replacing the pseudotree with the descriptive unravelling could give way to a similar result for alternative Vietoris-like coalgebras on Stone spaces.

In [19, 20], the van Benthem characterisation Theorem is treated for intuitionistic frames. Descriptive intuitionistic frames are known as Esakia spaces

[9] and with the techniques from this paper, one could pursue modal characterisation theorems on these interesting classes.

Finally, neighbourhood structures have been given the structure of Stone coalgebras in [13]. Thus, the constructions presented here may be combined with the approach from [12] to achieve a modal characterisation result on these neighbourhood structures over Stone spaces.

Acknowledgements The authors are very thankful to Ian Hodkinson for many interesting discussions (some of which date back to late 2000's when the first-named author was a research fellow at Imperial College London), valuable suggestions that substantially improved the manuscript and for spotting a gap in the original version of the main proof.

References

[1] J. v. Benthem. *Modal correspondence theory.* PhD thesis, University of Amsterdam, 1976.

[2] N. Bezhanishvili, G. Fontaine, and Y. Venema. Vietoris bisimulations. *Journal of Logic and Computation*, 20(5):1017–1040, 2010.

[3] N. Bezhanishvili and I. Hodkinson. Sahlqvist theorem for modal fixed point logic. *Theoret. Comput. Sci.*, 424:1–19, 2012.

[4] P. Blackburn, M. De Rijke, and Y. Venema. Modal logic (cambridge tracts in theoretical computer science), 2002.

[5] A. Chagrov and M. Zakharyaschev. *Modal logic*, volume 35 of *Oxford Logic Guides*. The Clarendon Press, Oxford University Press, New York, 1997. Oxford Science Publications.

[6] A. Dawar and M. Otto. Modal characterisation theorems over special classes of frames. *Annals of Pure and Applied Logic*, 161(1):1–42, 2009.

[7] H.-D. Ebbinghaus and J. Flum. *Finite model theory.* Springer Science & Business Media, 2005.

[8] S. Enqvist and S. Sourabh. Bisimulations for coalgebras on Stone spaces. *Journal of Logic and Computation*, 28(6):991–1010, 2018.

[9] L. L. Esakia. Topological Kripke models. *Dokl. Akad. Nauk SSSR*, 214:298–301, 1974.

[10] E. Grädel, P. G. Kolaitis, L. Libkin, M. Marx, J. Spencer, M. Y. Vardi, Y. Venema, and S. Weinstein. *Finite Model Theory and its applications.* Springer Science & Business Media, 2007.

[11] W. Hanf. Model-theoretic methods in the study of elementary logic. In *The theory of models*, pages 132–145. Elsevier, 2014.

[12] H. Hansen, C. Kupke, and E. Pacuit. Neighbourhood Structures: Bisimilarity and Basic Model Theory. *Logical Methods in Computer Science*, 5(2), 2009.

[13] H. H. Hansen and C. Kupke. A coalgebraic perspective on monotone modal logic. *Electronic Notes in Theoretical Computer Science*, 106:121–143, 2004.

[14] T. Henke. The van Benthem Characterisation Theorem for Descriptive Models. Master's thesis, Universiteit van Amsterdam, July 2019.

[15] D. Janin and I. Walukiewicz. On the expressive completeness of the propositional mu-calculus with respect to monadic second order logic. In *International Conference on Concurrency Theory*, pages 263–277. Springer, 1996.

[16] B. Jónsson and A. Tarski. Boolean algebras with operators. Part I. *American journal of mathematics*, 73(4):891–939, 1951.

[17] M. Kracht. *Tools and techniques in modal logic*, volume 97. Elsevier Amsterdam, 1999.

[18] C. Kupke, A. Kurz, and Y. Venema. Stone coalgebras. *Theoretical Computer Science*, 327(1-2):109–134, 2004.

[19] G. Olkhovikov. Model-theoretic characterization of intuitionistic predicate formulas. *J. Log. Comput.*, 24(4):809–829, 2014.

[20] G. Olkhovikov. On generalized Van Benthem-type characterizations. *Annals of Pure and Applied Logic*, 168(9):1643–1691, 2017.

[21] E. Rosen. Modal logic over finite structures. *Journal of Logic, Language and Information*, 6(4):427–439, 1997.

[22] L. Schröder, D. Pattinson, and T. Litak. A van Benthem/Rosen theorem for coalgebraic predicate logic. *Journal of Logic and Computation*, 27(3):749–773, 2017.

[23] Y. Venema. Algebras and coalgebras. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in logic and practical reasoning*, pages 331–426. North-Holland, 2007.

[24] Y. Venema and J. Vosmaer. Modal logic and the Vietoris functor. In *Leo Esakia on Duality in Modal and Intuitionistic Logics*, pages 119–153. Springer, 2014.