Characterisations of two basic hybrid logics

Ian Hodkinson*

Department of Computing, Imperial College London, London SW7 2AZ, UK

i.hodkinson@imperial.ac.uk   www.doc.ic.ac.uk/~imh

May 29, 2025

Abstract

We prove a van Benthem–Rosen-style characterisation theorem for two basic hybrid logics: modal logic with nominals, and modal logic with nominals and @. In each case, we show that over all Kripke models, and over all finite Kripke models, every first-order formula that is invariant under the appropriate bisimulations is equivalent to a hybrid formula, and we give optimal bounds on its modal depth in terms of the quantifier depth of the first-order formula.

We also show that the characterisation for modal logic with nominals and @ extends to arbitrary bisimulation-closed classes of Kripke models and to the class of finite models within such classes, while the characterisation for modal logic with nominals alone does not.

Keywords: modal logic; nominals; van Benthem Rosen theorem; hybrid logic characterisation theorem; bisimulation; bisimulation-closed classes; finite models.

MSC2020 classification: primary 03B45; secondary 03C13, 68Q85, 91A44.

1 Introduction

A basic fact about Kripke semantics for modal logic is van Benthem's theorem [6, 7] that up to logical equivalence, modal logic 'is' the bisimulation-invariant fragment of first-order logic — as far as formulas $\varphi(x)$ with at most one free variable are concerned, and in a signature comprising only unary and binary relation symbols. Modal logic is thus expressively complete for this fragment, and provides an effective syntax for it (the fragment itself is undecidable). Van Benthem's proof used the compactness theorem for first-order logic, and it applies to every elementary class of Kripke models.

This 'modal characterisation theorem' has attracted enormous interest, and a vast number of extensions have been found. Two kinds of extension are directly relevant to this note. On the one hand, Rosen [19] extended van Benthem's 'classical' result to *finite models,* showing that every first-order formula $\varphi(x)$ that is bisimulation invariant over finite Kripke models is equivalent to a modal formula over finite models. This does not follow from the classical result because some first-order formulas are bisimulation invariant over finite models but not over all models [17, 18]. Since its conclusion is stronger, the classical result is not an immediate consequence of the result in the finite either. One might ask if it follows with some extra effort, but Rosen rendered this question moot by providing a

*Many thanks to Guillermo Badia and Rob Goldblatt for valuable comments and encouragement. I also offer sincere thanks to the referees for taking the time to read the paper and for their valuable advice, suggestions, and encouragement. For example, before their interventions, §8.2.2 was just one sentence! Finally I thank Jane Spurr and the Editors for handling the paper.

uniform argument for both the classical and finite cases, so reproving van Benthem's original result in a different way. Rosen's proof used Hanf locality rather than compactness, which fails in the finite. Otto [17, 18] gave an 'elementary' version of the proof, replacing Hanf locality by a direct application of Ehrenfeucht–Fraïssé games, and establishing an optimal bound $2^q - 1$ on the modal depth of an equivalent modal formula in terms of the quantifier depth q of $\varphi(x)$.

On the other hand, different notions of bisimulation have been given for various *hybrid logics,* and some characterisation theorems have been proved for them — see, e.g., [2, 10, 4, 15, 3, 5]. So far, results have been by and large classical, proved using compactness or ultraproducts, and do not cover finite models. But some do. For example, Abramsky and Marsden [1, theorem 11] characterised the temporal hybrid logic with $\downarrow$ and @ in terms of invariance under generated submodels and/or disjoint unions, again establishing the classical and finite cases uniformly. They state [1, §7] that the result still holds in the presence of nominals. Though not directly concerned with hybrid logic, [21] proves an immensely general coalgebraic characterisation theorem for a range of modal-like logics, again uniformly for all models and for finite models.

In this note, we prove a characterisation theorem for two basic hybrid logics, providing a uniform proof that works both classically and in the finite, as Rosen and Otto did, and giving optimal modal depth bounds as Otto did.

The first hybrid logic is simply modal logic with *nominals* — special propositional atoms that are true at precisely one point of each model. We could perhaps call it *'proto-hybrid logic'*. It is, to be sure, a minimal extension of modal logic, but still a 'far from negligible' one [9, p.49]. We use ordinary modal bisimulations here, showing that both classically and in the finite, proto-hybrid logic is the bisimulation-invariant fragment of first-order logic in signatures comprising constants as well as unary and binary relation symbols. Van Benthem's original theorem shows this in the classical case, since being a nominal is first-order definable and so the class of relevant models is elementary. But I have not found a characterisation theorem in the literature for proto-hybrid logic over finite models, nor any depth bounds.

The second hybrid logic is actually called *'basic hybrid logic'* [9, §6.2]. It adds the hybrid actuality operator @ to proto-hybrid logic. A classical characterisation theorem is known for this logic — see [2, theorem 6.1], recalled in [9, theorem 39] — but again, I am not aware of one for finite models, nor any depth bounds. The appropriate notion of bisimulation [9, definition 37] is now slightly stronger, and the bisimulation-invariant first-order fragment consequently slightly larger, but the difference is so slight that we can handle both basic hybrid and proto-hybrid logics here using much the same proof.

The proof itself follows standard lines. The key is to show that every bisimulation-invariant first-order formula $\varphi(x)$ is *'local'* — that is to say, invariant under passing to a 'local neighbourhood' that the hybrid logic can control. The neighbourhood is typically the set of points 'near' to x.

In a little more detail, we want to find a finite set $\mathcal{H}$ of hybrid formulas such that any two pointed Kripke models that agree on $\mathcal{H}$ also agree on φ. For φ will then be equivalent to a boolean combination of formulas in $\mathcal{H}$.

Locality, if we can establish it, lets us restrict each of the two models to a 'neighbourhood' without changing the value of φ. The two neighbourhoods should also agree on $\mathcal{H}$: and a sufficiently large $\mathcal{H}$ should control them well enough to ensure that they also agree on φ — for example, because they are bisimilar. We are done.

This trail was blazed by Rosen [19] and most later writers have followed it. What is perhaps novel here is the choice of neighbourhoods. *Modal unravellings* [9, §3.2] are often used to simplify neighbourhoods sufficiently for $\mathcal{H}$ to control. (Sometimes $\mathcal{H}$ is so powerful that this is unnecessary

[1].) Unfortunately, unravellings involve duplicating points in a model. This is problematic with nominals, which must remain true at only one point. So we will use fairly obvious but perhaps new unravellings able to handle nominals. Notwithstanding this, neighbourhoods are more complicated in the presence of nominals, and nominals also interfere to a degree with *disjoint unions,* an ingredient of the proof of locality. We will therefore *interpret* the unravellings (in the model-theoretic sense) in simpler and better-behaved models.

These changes are not wholly trivial, because for basic hybrid logic, the optimal bound on the depth of the equivalent hybrid formula is larger than Otto's bound for modal logic. For proto-hybrid logic, it is larger still.

There is another way in which hybrid logic behaves differently from modal logic. In [17], Otto gave a bisimulation characterisation for modal logic over arbitrary bisimulation-closed classes of Kripke models, both classically and in the finite. We will prove an analogous result for basic hybrid logic, but show that there is no such result for proto-hybrid logic.

I tried to prove the characterisation theorems in this note because I wanted to know whether they were true in the finite. This is not a given. As a warning, while van Benthem's theorem shows that classically, modal logic is the bisimulation-invariant fragment of first-order logic over transitive models, this fails in the finite and additional modal connectives are needed [11]. Another (non-modal) warning example is furnished by two-variable first-order logic [17, 18]. And for proto-hybrid logic over bisimulation-closed classes, there is no characterisation result classically or in the finite. More motivation for the results will be given in section 9 in the light of the greater context available at that point.

Layout. Sections 2–4 present background material and notation, increasingly specialised as we proceed, and with some slight novelties in §4.4 and §4.5, but overall with few surprises. Readers will most likely be familiar with this material, so the treatment is brief, but still it takes up around half the paper. Readers may of course skip it and refer back to it as needed. The real work begins in section 5, where we define and study the unravellings. The main characterisation theorem is in section 6. Section 7 gives examples to show optimality of modal depth bounds, and section 8 looks at possible extensions of the theorem to other logics and classes: in particular, bisimulation-closed classes. The conclusion in section 9 has some comments, such as on possible further work.

2 Definitions

This section presents basic definitions and notation, both for general matters and for the hybrid logics we consider.

2.1 Generalities

We use standard (von Neumann) ordinals. Each ordinal α is the set of smaller ordinals, so the smallest infinite ordinal ω is $\{0, 1, \ldots\}$, and $n = \{0, 1, \ldots, n-1\}$ and $n+1 = n \cup \{n\}$ for $n < \omega$. Ordinal sum $\alpha + \beta$ is defined as usual (as the order type of α followed by β) — for example, $1 + \omega = \omega < \omega + 1$. We write $|S|$ for the cardinality of a set S, and use $\wp$ to denote the power-set operation. We let $S \uplus T$ denote the disjoint union of sets S and T. It can be defined formally as $S \times \{0\} \cup T \times \{1\}$, but we treat it informally. We generally write binary relations in infix form.

We write id_S for the identity map or function on a set S. For sets S, T, we let ${}^S T$ denote the set $\{f \mid f : S \to T\}$ of functions from S to T. We write $\operatorname{dom} f$ and $\operatorname{rng} f$ for the domain and range, respectively, of a function f. For $S \subseteq \operatorname{dom} f$, we write $f \restriction S$ for the restriction of f to S, and $f(S)$ for $\{f(s) : s \in S\}$ (this can be ambiguous but causes us no difficulties). Similarly, for a family $\boldsymbol{a} = (a_i : i \in I) \in {}^I(\operatorname{dom} f)$, we write $f(\boldsymbol{a})$ for $(f(a_i) : i \in I) \in {}^I(\operatorname{rng} f)$.

2.2 Hybrid logic

For hybrid logic, we broadly follow the notation in [9]. A *hybrid signature* is a set σ partitioned into two sets, PROP and NOM, where PROP denotes the set of *propositional atoms* (or *propositional variables*) and NOM denotes the set of *nominals*. For a hybrid signature $\sigma = \text{PROP} \cup \text{NOM}$, we define the hybrid $\mathcal{L}^{\diamond @}(\sigma)$*-formulas* ψ, and their *modal depth* $d(\psi)$, as follows.

1. Each element of σ is an $\mathcal{L}^{\diamond @}(\sigma)$-formula, of modal depth 0.
2. $\top$ is an $\mathcal{L}^{\diamond @}(\sigma)$-formula, also of modal depth 0.
3. If ψ and θ are $\mathcal{L}^{\diamond @}(\sigma)$-formulas, then so are:
   (a) $\neg\psi$, and $d(\neg\psi) = d(\psi)$,
   (b) $\psi \wedge \theta$, and $d(\psi \wedge \theta) = \max(d(\psi), d(\theta))$,
   (c) $\diamond\psi$, and $d(\diamond\psi) = 1 + d(\psi)$.
4. If ψ is an $\mathcal{L}^{\diamond @}(\sigma)$-formula and $c \in \text{NOM}$, then $@_c\psi$ is an $\mathcal{L}^{\diamond @}(\sigma)$-formula, and $d(@_c\psi) = d(\psi)$.

An $\mathcal{L}^{\diamond}(\sigma)$*-formula* is a $\mathcal{L}^{\diamond @}(\sigma)$-formula that does not involve any @ — that is, we drop clause 4 above. So $\mathcal{L}^{\diamond}(\sigma)$-formulas are just modal formulas, except that they may involve nominals. $\mathcal{L}^{\diamond}$ and $\mathcal{L}^{\diamond @}$ are the languages of *proto-hybrid logic* and *basic hybrid logic,* respectively. We regard $\bot, \vee, \to, \leftrightarrow, \Box$ as the usual abbreviations. For a $\mathcal{L}^{\diamond @}$-formula ψ and $n < \omega$, we define $\diamond^n\psi$ by induction: $\diamond^0\psi = \psi$, and $\diamond^{n+1}\psi = \diamond\diamond^n\psi$. For a nonempty finite set $S = \{\psi_0, \ldots, \psi_{n-1}\}$ of $\mathcal{L}^{\diamond @}(\sigma)$-formulas, we write $\bigwedge S$ and $\bigwedge_{i<n} \psi_i$ for $\psi_0 \wedge \ldots \wedge \psi_{n-1}$, and $\bigvee S$ and $\bigvee_{i<n} \psi_i$ for $\psi_0 \vee \ldots \vee \psi_{n-1}$; the order and bracketing of the ψ_i is immaterial (semantically). We let $\bigwedge \emptyset = \top$ and $\bigvee \emptyset = \bot$.

A *Kripke model* (for σ) is a triple $M = (W, R^M, V)$, where $W \neq \emptyset$ is the set of 'worlds', also called the *domain* of M, $R^M \subseteq W \times W$ is the 'accessibility relation', and $V : \sigma \to \wp(W)$ is the 'valuation', satisfying $|V(c)| = 1$ for each $c \in \text{NOM}$ — we write c^M for the unique element of $V(c)$. We write $\operatorname{dom}(M)$, and more often just M, for its domain W. We say that M is *finite* if it has finite domain.

A *submodel of* M is a Kripke model of the form $N = (U, R^M \cap (U \times U), V_U)$, where $\emptyset \neq U \subseteq W$ and $V_U(p) = V(p) \cap U$ for $p \in \sigma$. (This is a well-defined Kripke model iff $c^M \in U$ for each $c \in \text{NOM}$.) We say that N is a *generated submodel of* M if $u \in U$, $w \in W$, and uR^Mw imply $w \in U$.

When we consider a hybrid signature $\tau \subseteq \sigma$, it will be implicit that the type (atom or nominal) of each symbol in τ is inherited from σ. We write $M \restriction \tau$ for the Kripke model $(W, R^M, V \restriction \tau)$ for τ, called the τ*-reduct of* M.

We define the semantics of $\mathcal{L}^{\diamond @}(\sigma)$-formulas in Kripke models $M = (W, R^M, V)$ for σ as usual: for $w \in W$, we define

1. $M, w \models p$ iff $w \in V(p)$, for $p \in \sigma$,
2. $M, w \models \top$,

3. $M, w \models \neg\psi$ iff $M, w \not\models \psi$,

4. $M, w \models \psi \wedge \theta$ iff $M, w \models \psi$ and $M, w \models \theta$,

5. $M, w \models \Diamond\psi$ iff $M, u \models \psi$ for some $u \in W$ with $wR^M u$,

6. $M, w \models @_c\psi$ iff $M, c^M \models \psi$.

A *pointed Kripke model* (for σ) is a pair (M, w), where M is a Kripke model (for σ) and $w \in M$. We say that (M, w) is *finite* if M is finite. We say that pointed Kripke models (A, a) and (B, b) for σ *agree* on an $\mathcal{L}^{\Diamond @}(\sigma)$-formula ψ if $A, a \models \psi \iff B, b \models \psi$, and *agree* on a set S of $\mathcal{L}^{\Diamond @}(\sigma)$-formulas if they agree on every formula in S. 'Disagree' will mean 'do not agree'.

3 Bisimulations and games

Fix, for this section, a hybrid signature $\sigma = \text{PROP} \cup \text{NOM}$. All Kripke models in this section are for σ.

3.1 Bisimulations

Much of this note is concerned with bisimulations. A $\Diamond$-*bisimulation* (generally called just a *bisimulation* in the literature) between Kripke models A and B is a binary relation $Z \subseteq A \times B$ such that for each $a \in A$ and $b \in B$ with aZb,

1. (A, a) and (B, b) agree on σ,

2. ('Forth') if $a' \in A$ and $aR^A a'$, then there is $b' \in B$ with $bR^B b'$ and $a'Zb'$,

3. ('Back') if $b' \in B$ and $bR^B b'$, then there is $a' \in A$ with $aR^A a'$ and $a'Zb'$.

Z is said to be a $\Diamond @$-*bisimulation* (called a 'bisimulation-with-names' in [9, §6.2]) if it also satisfies:

4. $c^A Z c^B$ for each $c \in \text{NOM}$.

The difference is that a plain $\Diamond$-bisimulation may not relate c^A to anything, nor c^B. It goes without saying that every $\Diamond @$-bisimulation is a $\Diamond$-bisimulation.

DEFINITION 3.1 For $\star \in \{\Diamond, \Diamond @\}$, we say that pointed Kripke models (A, a) and (B, b) are $\star$-*bisimilar,* and write $(A, a) \sim^\star (B, b)$, if there is a $\star$-bisimulation Z between A and B such that aZb.

EXAMPLE 3.2 If A, B are Kripke models and A is a generated submodel of B, then the inclusion map $\iota : A \hookrightarrow B$ is a $\Diamond @$-bisimulation, so $(A, a) \sim^{\Diamond @} (B, a)$ for every $a \in A$.

Perhaps we should say rather that the *graph* $\{(a, \iota(a)) : a \in A\}$ of ι is a bisimulation, but set-theoretically, a function is its graph.

FACT 3.3 For each $\star \in \{\Diamond, \Diamond @\}$, $\mathcal{L}^\star(\sigma)$-formulas are $\star$-*bisimulation invariant:* ie. if (A, a) and (B, b) are pointed Kripke models and $(A, a) \sim^\star (B, b)$, then (A, a) and (B, b) agree on all $\mathcal{L}^\star(\sigma)$-formulas. See, e.g., [9, lemmas 9 and 38].

3.2 Games

Bisimulations can be simulated by games. Let (A, a) and (B, b) be pointed Kripke models and let $\alpha \leq \omega$ be an ordinal. We define an α-round game $\mathrm{Bs}^{\Diamond}_{\alpha}(A, a, B, b)$ as follows. There are two players, $\forall$ and $\exists$. The successive rounds are numbered $0, 1, \ldots, t, \ldots$ for $t < \alpha$. The initial position, regarded as chosen by $\forall$, is defined by $a_0 = a$ and $b_0 = b$, and $\exists$ loses outright, before any rounds are played and even if $\alpha = 0$, if (A, a_0) and (B, b_0) disagree on σ. By the way, games such as $\mathrm{Bs}^{\Diamond}_0$, with no rounds, are well defined and indeed useful — see lemma 3.7, for example.

At the start of each round $t < \alpha$, points $a_t \in A$ and $b_t \in B$ are already chosen. In the round, $\forall$ chooses some $a_{t+1} \in A$ with $a_t R^A a_{t+1}$, or some $b_{t+1} \in B$ with $b_t R^B b_{t+1}$. He loses if he can't do this. With full knowledge of his move, $\exists$ must respond with some $b_{t+1} \in B$ with $b_t R^B b_{t+1}$, or some $a_{t+1} \in A$ with $a_t R^A a_{t+1}$, respectively, and she loses if she can't. That completes the round, and $\exists$ loses the game at this point if (A, a_{t+1}) and (B, b_{t+1}) disagree on σ. $\exists$ wins if she never loses at any stage.

The game $\mathrm{Bs}^{\Diamond @}_{\alpha}(A, a, B, b)$ is the same, except that $\forall$ is allowed to choose the initial position (a_0, b_0) to be any pair in the set $\{(a, b), (c^A, c^B) : c \in \text{NOM}\}$.

A *strategy for* $\exists$ in any of the games in this note is a set of rules telling $\exists$ how to move in any position that can arise when she uses the strategy. A strategy is said to be *winning* if $\exists$ wins every play of the game in which she uses it.

DEFINITION 3.4 Let (A, a), (B, b) be pointed Kripke models, $\star \in \{\Diamond, \Diamond @\}$, and $\alpha \leq \omega$. We write $(A, a) \sim^{\star}_{\alpha} (B, b)$ if $\exists$ has a winning strategy in the game $\mathrm{Bs}^{\star}_{\alpha}(A, a, B, b)$.

The following is an elementary games lemma.

LEMMA 3.5 *Let (A, a), (B, b) be pointed Kripke models, $\star \in \{\Diamond, \Diamond @\}$, and $\alpha \leq \omega$.*

1. $(A, a) \sim^{\Diamond @}_{\alpha} (B, b)$ *iff* $(A, a') \sim^{\star}_{\alpha} (B, b')$ *for every* $(a', b') \in \{(a, b), (c^A, c^B) : c \in \text{NOM}\}$.
2. *If* $(A, a) \sim^{\star}_{\alpha} (B, b)$ *then* $(A, a) \sim^{\star}_{\beta} (B, b)$ *for every* $\beta < \alpha$.
3. *If* $(A, a) \sim^{\star}_{1+\alpha} (B, b)$ *then the following all hold:*
   (a) $(A, a) \sim^{\star}_{0} (B, b)$,
   (b) *for each* $a' \in A$ *with* $aR^A a'$, *there is some* $b' \in B$ *with* $bR^B b'$ *and* $(A, a') \sim^{\star}_{\alpha} (B, b')$,
   (c) *for each* $b' \in B$ *with* $bR^B b'$, *there is some* $a' \in A$ *with* $aR^A a'$ *and* $(A, a') \sim^{\star}_{\alpha} (B, b')$.

   The converse implication holds when $\star = \Diamond$.
4. $(A, a) \sim^{\star}_{\omega} (B, b)$ *iff* $(A, a) \sim^{\star} (B, b)$.

3.3 Games and formulas

In this subsection, we assume that $\sigma = \text{PROP} \cup \text{NOM}$ is finite. This assumption is used in the following definition to ensure that the $\mathcal{F}_k$ are sets of formulas, and in lemma 3.7($3 \Rightarrow 1$).

DEFINITION 3.6 We define, by induction on $k < \omega$, a finite set $\mathcal{F}_k$ (also written $\mathcal{F}^{\Diamond}_k$) of $\mathcal{L}^{\Diamond}(\sigma)$-formulas, and a finite set $\mathcal{F}^{\Diamond @}_k$ of $\mathcal{L}^{\Diamond @}(\sigma)$-formulas, as follows:

$$\begin{aligned} \mathcal{F}_0 &= \sigma, \\ \mathcal{F}_{k+1} &= \sigma \cup \{\Diamond(\textstyle\bigwedge S \wedge \neg \bigvee(\mathcal{F}_k \setminus S)) : S \subseteq \mathcal{F}_k\}, \\ \mathcal{F}^{\Diamond @}_k &= \mathcal{F}_k \cup \{@_c \psi : c \in \text{NOM},\ \psi \in \mathcal{F}_k\}. \end{aligned}$$

The proof of the following lemma is quite standard, but we include a sketch to illustrate the games and show how @ is handled.

LEMMA 3.7 *Assuming σ finite, let $(A, a), (B, b)$ be pointed Kripke models, $\star \in \{\Diamond, \Diamond@\}$, and $k < \omega$. The following are equivalent:*

1. $(A, a) \sim_k^\star (B, b)$,
2. (A, a) *and* (B, b) *agree on all* $\mathcal{L}^\star(\sigma)$*-formulas of modal depth* $\leq k$,
3. (A, a) *and* (B, b) *agree on* $\mathcal{F}_k^\star$.

Proof. For $1 \Rightarrow 2$, by lemma 3.5(2) it suffices to prove by induction on $\mathcal{L}^\star(\sigma)$-formulas ψ that if $(A, a) \sim_{d(\psi)}^\star (B, b)$ then $A, a \models \psi$ iff $B, b \models \psi$. For $\psi \in \sigma \cup \{\top\}$ this is clear. Assume the result for ψ and θ inductively. The case $\neg\psi$ is very simple, the case $\psi \wedge \theta$ follows from lemma 3.5(2), and the case $\Diamond\psi$ from lemma 3.5(3). For the case $@_c\psi$ for a nominal c, suppose that $(A, a) \sim_{d(@_c\psi)}^{\Diamond@} (B, b)$ (the case $\star = \Diamond$ is of course impossible here). By lemma 3.5(1) and because $d(@_c\psi) = d(\psi)$, we have $(A, c^A) \sim_{d(\psi)}^{\Diamond@} (B, c^B)$, so inductively, $A, c^A \models \psi$ iff $B, c^B \models \psi$. By definition of the semantics of @ in §2.2, we obtain $A, a \models @_c\psi$ iff $B, b \models @_c\psi$, as required.

Part 3 follows from part 2 since all formulas in $\mathcal{F}_k^\star$ have modal depth $\leq k$.

We first prove $3 \Rightarrow 1$ for $\star = \Diamond$. For a pointed Kripke model (M, m), write $\mathrm{tp}_k(M, m) = \{\psi \in \mathcal{F}_k : M, m \models \psi\}$. Then for $S \subseteq \mathcal{F}_k$ we have

$$\mathrm{tp}_k(M, m) = S \quad \text{iff} \quad M, m \models \bigwedge S \wedge \neg \bigvee(\mathcal{F}_k \setminus S). \tag{1}$$

We now show by induction on k that if (A, a) and (B, b) agree on $\mathcal{F}_k$ then $(A, a) \sim_k^\Diamond (B, b)$. For $k = 0$ it's clear. Assume the result for k and suppose that (A, a) and (B, b) agree on $\mathcal{F}_{k+1}$. We establish (a)–(c) of lemma 3.5(3). For (a), certainly $(A, a) \sim_0^\Diamond (B, b)$ since $\sigma \subseteq \mathcal{F}_{k+1}$. For (b), take any $a' \in A$ with aR^Aa', and let $S = \mathrm{tp}_k(A, a')$ and $\chi = \bigwedge S \wedge \neg \bigvee(\mathcal{F}_k \setminus S)$. By (1), $A, a' \models \chi$, so by semantics of $\Diamond$ we get $A, a \models \Diamond\chi$. This formula is in $\mathcal{F}_{k+1}$, so $B, b \models \Diamond\chi$ as well. Again by semantics of $\Diamond$, there is $b' \in B$ with bR^Bb' and $B, b' \models \chi$ — and (1) gives $\mathrm{tp}_k(B, b') = S = \mathrm{tp}_k(A, a')$. So (A, a') and (B, b') agree on $\mathcal{F}_k$. Inductively, $(A, a') \sim_k^\Diamond (B, b')$. Similarly, we can prove that (c) for every $b' \in B$ with bR^Bb', there is $a' \in A$ with aR^Aa' and $(A, a') \sim_k^\Diamond (B, b')$. So by the 'converse implication' of lemma 3.5(3), $(A, a) \sim_{k+1}^\Diamond (B, b)$. This completes the induction and proves $3 \Rightarrow 1$ for $\star = \Diamond$.

Finally suppose that (A, a) and (B, b) agree on $\mathcal{F}_k^{\Diamond@}$. By definition of $\mathcal{F}_k^{\Diamond@}$ and semantics of @, (A, a') and (B, b') agree on $\mathcal{F}_k$ for every $(a', b') \in \{(a, b), (c^A, c^B) : c \in \textsc{Nom}\}$. By the $\Diamond$-case above, $(A, a') \sim_k^\Diamond (B, b')$ for every $(a', b') \in \{(a, b), (c^A, c^B) : c \in \textsc{Nom}\}$. By lemma 3.5(1), $(A, a) \sim_k^{\Diamond@} (B, b)$, as required. □

4 Classical logics

The purpose of this note is to compare basic hybrid logic with classical first-order logic, so we discuss the latter now. In fact, we go via *infinitary logic,* which we will use in interpretations below. Infinitary logic has been a basic part of model theory since the 1960s — Hodges' model theory text [14, §2.1] introduces it even before first-order logic — and I make no apologies for using it here. It is actually needed in only one place, in definition 5.6(2), and then only when the structure A is infinite. In the finite-models case, all 'infinitary' formulas in the paper are actually first-order.

See [14] for more information on the topics below.

4.1 Classical infinitary logic

The following is standard and we include it mainly to fix names and notation. A (classical) *signature* is a set L of relation symbols with specified finite arities, and constants. *In this note, we do not need function symbols and will not consider them.* We say that L is *relational* if it contains no constants.

The $L_{\infty\omega}$*-formulas* φ, together with the (perhaps infinite) set $FV(\varphi)$ of free variables of φ and the (ordinal) *quantifier depth of* φ, are defined as in first-order logic with equality but allowing conjunctions and disjunctions over arbitrary sets of formulas. See, e.g., [14, §2.1]. Nearly all formulas that we consider will have finite quantifier depth. An $L_{\infty\omega}$-formula is said to be *atomic* if it has no proper subformulas, *quantifier-free* if it has no quantifiers, and *first-order,* or just an L*-formula,* if every conjunction and disjunction in it is over a finite set.

An L*-structure* M comprises a nonempty set $\mathrm{dom}(M)$, the *domain* of M, together with an *interpretation* s^M of each $s \in L$ as an n-ary relation on $\mathrm{dom}(M)$, if s is an n-ary relation symbol, and an element of $\mathrm{dom}(M)$ if s is a constant. As with Kripke models, we usually write $\mathrm{dom}(M)$ as just M. We say that M is *finite* if its domain is.

For an L-structure M and an 'assignment' h mapping variables into $\mathrm{dom}(M)$, we define $M, h \models \varphi$ for each $L_{\infty\omega}$-formula φ in the usual way. For a formula φ, an index set I, and pairwise distinct variables x_i $(i \in I)$, we write $\varphi(x_i : i \in I)$ to indicate that $FV(\varphi) \subseteq \{x_i : i \in I\}$. We sometimes stretch this notation slightly, writing, e.g., $\varphi(x_1, \ldots, x_n, (v_i : i \in I))$ to indicate that $FV(\varphi) \subseteq \{x_1, \ldots, x_n\} \cup \{v_i : i \in I\}$. As usual, whether $M, h \models \varphi$ or not depends only on $h \restriction FV(\varphi)$ (and on M and φ of course). So for a formula $\varphi(x_i : i \in I)$ and elements $a_i \in M$ $(i \in I)$, we can write $M \models \varphi(a_i : i \in I)$ if $M, h \models \varphi$, where $h(x_i) = a_i$ for each $i \in I$.

A *substructure* of M is an L-structure N with $\mathrm{dom}(N) \subseteq \mathrm{dom}(M)$ and $N \models \alpha(a_1, \ldots, a_n)$ iff $M \models \alpha(a_1, \ldots, a_n)$ for each atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in N$. The latter holds iff $R^N = R^M \cap \mathrm{dom}(N)^n$ for each n-ary relation symbol $R \in L$, and $c^M \in N$ and $c^M = c^N$ for each constant $c \in L$.

Let M, N be L-structures. A map $f : M \to N$ is said to be an *(L)-homomorphism* if for every atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in M$, we have $M \models \alpha(a_1, \ldots, a_n) \implies N \models \alpha(f(a_1), \ldots, f(a_n))$. A partial map $f : M \to N$ is said to be an *(L)-partial isomorphism* if $M \models \alpha(a_1, \ldots, a_n) \iff N \models \alpha(f(a_1), \ldots, f(a_n))$ for every atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in \mathrm{dom}\, f$; if also $\mathrm{dom}\, f = M$, then f is called an *(L)-embedding;* and if also $M = N$ and $\mathrm{dom}\, f = \mathrm{rng}\, f = M$ then f is called an *automorphism of* M.

If L is relational, the *disjoint union* $A + B$ of L-structures A, B is the L-structure M defined informally by $\mathrm{dom}\, M = \mathrm{dom}\, A \uplus \mathrm{dom}\, B$ and $R^M = R^A \uplus R^B$ for each relation symbol $R \in L$. It is finite if A, B are finite. The inclusion maps $\ell : A \hookrightarrow M$ and $r : B \hookrightarrow M$ are L-embeddings.

4.2 Correspondence

Central to this note is the correspondence between hybrid and classical logic. For a hybrid signature $\sigma = \textsc{Prop} \uplus \textsc{Nom}$, the classical 'correspondence' signature $L(\sigma)$ comprises a binary relation symbol R; a unary relation symbol P for each $p \in \textsc{Prop}$; and the elements of $\textsc{Nom}$, taken as constants.

A Kripke model $M = (W, R^M, V)$ for σ can be viewed as an $L(\sigma)$-structure as follows. The domain of this structure is W. We interpret R as the binary relation R^M on W, we interpret P (for $p \in \textsc{Prop}$) as the unary relation $V(p)$ on W, and we interpret c (for $c \in \textsc{Nom}$) as c^M as defined in §2.2. We denote the resulting $L(\sigma)$-structure also by M. Conversely, an $L(\sigma)$-structure can be construed as a Kripke model for σ in the obvious way. So we will regard a Kripke model for σ equally as an $L(\sigma)$-structure, making no distinction between them.

Let σ be a hybrid signature and $\mathcal{C}$ a class of pointed Kripke models for σ. Let $\varphi(x)$ be an $L(\sigma)$-formula, and ψ an $\mathcal{L}^{\Diamond @}(\sigma)$-formula. We say that φ and ψ are *equivalent over* $\mathcal{C}$ if $M \models \varphi(w)$ iff $M, w \models \psi$, for every $(M, w) \in \mathcal{C}$.

We can also in a sense view $\mathcal{L}^{\Diamond @}(\sigma)$-formulas as first-order $L(\sigma)$-formulas, via their *standard translations.* We will not use these (see, e.g., [2, proposition 3.1] and [4, proposition 11] for the definition), but to give the idea, for $c, d \in \textsc{Nom}$ and $p \in \textsc{Prop}$, the standard translation of $\Diamond(c \wedge @_d p)$ is $\exists y(xRy \wedge y = c \wedge P(d))$. The standard translation of each $\mathcal{L}^{\Diamond @}(\sigma)$-formula ψ is an $L(\sigma)$-formula $\varphi(x)$ that is equivalent to ψ over every $\mathcal{C}$.

The converse question asks, for given $\mathcal{C}$ and $\star \in \{\Diamond, \Diamond @\}$, whether every $L(\sigma)$-formula $\varphi(x)$ is equivalent to some $\mathcal{L}^{\star}(\sigma)$-formula ψ over $\mathcal{C}$. By fact 3.3, $\mathcal{L}^{\star}(\sigma)$-formulas are $\star$-bisimulation invariant, so we restrict the question to those $\varphi(x)$ that are themselves *$\star$-bisimulation invariant over $\mathcal{C}$:* that is, $A \models \varphi(a)$ iff $B \models \varphi(b)$ whenever $(A, a), (B, b) \in \mathcal{C}$ and $(A, a) \sim^{\star} (B, b)$.

Assuming that $\varphi(x)$ is $\star$-bisimulation invariant over $\mathcal{C}$, we will answer the question affirmatively in theorem 6.1 for $\mathcal{C}$ the class of all pointed Kripke models for σ, and the class of finite ones; and in theorem 8.3 for $\star = \Diamond @$ and $\mathcal{C}$ any bisimulation-closed class of pointed Kripke models for σ (see definition 8.1) or the class of finite models in such a class. The latter result fails when $\star = \Diamond$ (example 8.2).

The following lemma, showing robustness of bisimulation invariance, will be helpful in theorem 6.1. The (easy) converse also holds, but we will not need it. There is no obvious analogue for bisimulation-closed classes: see example 8.4.

LEMMA 4.1 *Let $\tau \subseteq \sigma$ be hybrid signatures, let $\mathcal{C}_\tau$ and $\mathcal{C}_\sigma$ be the classes of all (or all finite) pointed Kripke models for τ and σ, respectively, let $\varphi(x)$ be an $L(\tau)$-formula, and $\star \in \{\Diamond, \Diamond @\}$. If φ is $\star$-bisimulation invariant over $\mathcal{C}_\sigma$, then it is also $\star$-bisimulation invariant over $\mathcal{C}_\tau$.*

Proof. Suppose $(A, a), (B, b) \in \mathcal{C}_\tau$ and $(A, a) \sim^{\star} (B, b)$. We show that $A \models \varphi(a)$ iff $B \models \varphi(b)$.[1]

Write $\sigma = \textsc{Prop} \uplus \textsc{Nom}$. The Kripke model A is for τ. Let A_1 be the Kripke model for σ defined by $A_1 \upharpoonright \tau = A$ (see §2.2 for the notation), $P^{A_1} = \emptyset$ for each $p \in \textsc{Prop} \setminus \tau$, and $c^{A_1} = a$ for each $c \in \textsc{Nom} \setminus \tau$. Plainly, $(A_1, a) \in \mathcal{C}_\sigma$, and $A \models \varphi(a)$ iff $A_1 \models \varphi(a)$ because A and A_1 agree on symbols in φ.

Define a second Kripke model A_2 for σ by adding to A_1 a new world $w \notin A \cup B$. Define each symbol in $L(\sigma)$ to have the exact same interpretation in A_2 as it does in A_1 (so w is an isolated world unrelated by R to any world). Then $(A_2, a) \in \mathcal{C}_\sigma$ as well. Plainly, A_1 is a generated submodel of A_2, so by example 3.2, the inclusion map $\iota : A_1 \hookrightarrow A_2$ is a $\star$-bisimulation. Since φ is assumed $\star$-bisimulation invariant over $\mathcal{C}_\sigma$, we obtain $A_1 \models \varphi(a)$ iff $A_2 \models \varphi(a)$.

Finally define a third Kripke model A_3 for σ. It is the same as A_2 except that each nominal in $\textsc{Nom} \setminus \tau$ is now interpreted as w. Then $A_2 \models \varphi(a)$ iff $A_3 \models \varphi(a)$, again because the two models agree on symbols in φ. Also, $(A_3, a) \in \mathcal{C}_\sigma$. Combining the three stages, we see that $A \models \varphi(a)$ iff $A_3 \models \varphi(a)$.

Now do the same for B, arriving at $(B_3, b) \in \mathcal{C}_\sigma$ with $B \models \varphi(b)$ iff $B_3 \models \varphi(b)$.

By assumption, $(A, a) \sim^{\star} (B, b)$. Let Z be a $\star$-bisimulation between A and B with aZb. It can be checked that $Z \cup \{(w, w)\}$ is a $\star$-bisimulation between A_3 and B_3. Since these models are in $\mathcal{C}_\sigma$,

[1] This is surprisingly tricky, mainly because σ might have nominals when τ does not. As a warning example, take $\tau = \emptyset \subseteq \sigma = \{c\}$, where c is a nominal, and consider the Kripke models $A = (\{0,1\}, \{(0,1),(1,1)\}, \emptyset)$ and $B = (\{2\}, \{(2,2)\}, \emptyset)$ for τ. Then $\{0,1\} \times \{2\}$ is a $\star$-bisimulation from A to B, so $(A, 0) \sim^{\star} (B, 2)$. But there do not exist Kripke models A', B' for σ with τ-reducts A, B (resp.) and with $(A', 0) \sim^{\star} (B', 2)$, because A' would have to make c true at both 0 and 1.

over which φ is assumed $\star$-bisimulation invariant, we obtain $A_3 \models \varphi(a)$ iff $B_3 \models \varphi(b)$. Putting all the steps together shows that $A \models \varphi(a)$ iff $B \models \varphi(b)$, and proves the lemma. □

4.3 Ehrenfeucht–Fraïssé games

Let L be a signature, let A, B be L-structures, let I be a possibly infinite index set, and let $a_i \in A$ and $b_i \in B$ for each $i \in I$. Write $\boldsymbol{a} = (a_i : i \in I)$ and $\boldsymbol{b} = (b_i : i \in I)$. (When I is a singleton $\{i\}$, we write $\boldsymbol{a}$ as simply a_i.) Let $q < \omega$ and suppose that $I \cap q = \emptyset$, to make things below well defined. The q-round Ehrenfeucht–Fraïssé game $\mathrm{EF}_q(A, \boldsymbol{a}, B, \boldsymbol{b})$ is played again by our players $\forall$ and $\exists$. The successive rounds are numbered $0, 1, \ldots, q-1$. In each round $t < q$, $\forall$ chooses a 'left element' $a_t \in A$, or a 'right element' $b_t \in B$.[2] Having seen $\forall$'s move, $\exists$ responds by choosing a right element $b_t \in B$ or a left element $a_t \in A$, respectively. That completes the round. At the end of play, $\exists$ wins if

$$A \models \alpha(a_{i_1}, \ldots, a_{i_n}) \quad \text{iff} \quad B \models \alpha(b_{i_1}, \ldots, b_{i_n}),$$

for every atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $i_1, \ldots, i_n \in I \cup q$.

DEFINITION 4.2 For $q < \omega$, we write $(A, \boldsymbol{a}) \equiv_{\infty q} (B, \boldsymbol{b})$ if $A \models \varphi(\boldsymbol{a})$ iff $B \models \varphi(\boldsymbol{b})$ for every $L_{\infty\omega}$-formula $\varphi(x_i : i \in I)$ of quantifier depth $\leq q$.

LEMMA 4.3 *If $\exists$ has a winning strategy in $\mathrm{EF}_q(A, \boldsymbol{a}, B, \boldsymbol{b})$, then $(A, \boldsymbol{a}) \equiv_{\infty q} (B, \boldsymbol{b})$.*

Proof. A standard exercise by induction on φ (like $1 \Rightarrow 2$ in lemma 3.7); or see, e.g., the proof of [14, theorem 3.5.2]. □

The converse of the lemma also holds, but we will not need it.

4.4 Weighted Gaifman graph

We extend the classical definition of *Gaifman graph* [12], as follows.[3]

DEFINITION 4.4 Let L be a signature, M an L-structure, C a substructure of M, and f an automorphism of C.

1. We define the graph $\mathfrak{G}(M, f)$ to be the undirected loopfree weighted graph with $\mathrm{dom}(M)$ as its set of nodes, and with the following edges and weights. Let $a, b \in M$ with $a \neq b$. Then:
   (a) if $a, b \in C$, and $b = f(a)$ or $a = f(b)$, then ab is an edge of weight 0.
   (b) Otherwise, if there are n-ary $R \in L$ and $a_1, \ldots, a_n \in M$ with $M \models R(a_1, \ldots, a_n)$ and $a, b \in \{a_1, \ldots, a_n\}$, then ab is an edge of weight 1.
   (c) Otherwise, ab is not an edge.

   We do not need to write $\mathfrak{G}(M, C, f)$, since we can recover C from M and f as the substructure of M with domain $\mathrm{dom}\, f = \mathrm{rng}\, f$. The identity map id_M on M is plainly an automorphism of M, and we write $\mathfrak{G}(M, id_M)$ as simply $\mathfrak{G}(M)$.

[2] We need this 'left–right' nomenclature in case the game has the form $\mathrm{EF}_q(A, \boldsymbol{a}, A, \boldsymbol{b})$. We could rename the second A as B, but in proposition 4.5 we do not want to do this.

[3] The extension will ensure that conditions R1–R4 in the proof of proposition 4.5 below are left–right symmetric.

2. We define distance $d^{M,f}$ in $\mathfrak{G}(M,f)$ by least-weight paths in the usual way. Formally, write $w(a,b)$ for the weight of an edge ab in $\mathfrak{G}(M,f)$. Then for each $a,b \in M$,

$$d^{M,f}(a,b) = \min\big\{ \textstyle\sum_{i<n} w(a_i,a_{i+1}) : n<\omega,\ a_0,\ldots,a_n \in M,\ a_0 = a, \\ a_n = b,\ a_i a_{i+1} \text{ is an edge of } \mathfrak{G}(M,f) \text{ for each } i<n\big\},$$

where we define empty sums to be zero (for when $n = 0$ and $a = b$) and $\min(\emptyset) = \infty$ (for when no such $a_0, \ldots, a_n$ exist).

3. For a family $\boldsymbol{a} = (a_i : i \in I)$ of elements of M, and $l < \omega$, the *open* (M,f)*-neighbourhood* of $\boldsymbol{a}$ of radius l is defined to be

$$\mathcal{N}_l^{M,f}(\boldsymbol{a}) = \{a \in M : d^{M,f}(a,a_i) < l \text{ for some } i \in I\}.$$

If $f = id_M$, we write this neighbourhood as simply $\mathcal{N}_l^M(\boldsymbol{a})$.

The function $d = d^{M,f} : M \times M \to \omega \cup \{\infty\}$ is an *extended pseudo-metric* on M — it satisfies, for each $a,b,c \in M$, the axioms $d(a,a) = 0$, $d(a,b) = d(b,a)$, and $d(a,c) \leq d(a,b) + d(b,c)$ (the *triangle inequality*), where $\leq$ and $+$ are extended from ω to $\omega \cup \{\infty\}$ in the usual way.

One might think of $\mathfrak{G}(M,f)$ and $d^{M,f}$ as the classical Gaifman graph and metric of the structure obtained by contracting each orbit of f to a single point. Conversely, we can recover the classical Gaifman graph of M by dropping the weights from $\mathfrak{G}(M)$. The function d^{M,id_M} is the usual Gaifman metric on M, and $\mathcal{N}_l^M(\boldsymbol{a})$ is the usual (open) Gaifman neighbourhood of $\boldsymbol{a}$ in M.

4.5 Locality

A key step in the modal characterisation theorems of Rosen and Otto was to show that every bisimulation-invariant first-order formula $\varphi(x)$ is *local:* invariant under restricting a model to a 'local neighbourhood' that, under the right circumstances, the modal logic can control. See [19, lemma 4], [17, theorem 3.1 step 1], [18, lemma 3.5], and [13, lemma 58]. The proofs were model-theoretic, via Hanf locality (Rosen) or Ehrenfeucht–Fraïssé games (Otto), and the method continues to be used to the present day — e.g., [21, theorem 27] and [1, 'workspace' lemma 13].

Proposition 4.5 and its corollary 4.6 below are close relatives of these results. They incorporate aspects of Rosen's and Otto's work and the proofs are quite simple. The desired locality will follow from corollary 4.6, but not as directly as in the cited references. We will obtain it in §5.3 via interpretations, to be discussed in §4.6.

PROPOSITION 4.5 *Let L be a signature, which as usual may contain constants and relation symbols but not function symbols. Let M be an L-structure, C a substructure of M, f an automorphism of C, and $\boldsymbol{a} = (a_i : i \in I)$ a family of elements of C. Let $q < \omega$, and assume that $\mathcal{N}_{2^q}^{M,f}(\boldsymbol{a}) \subseteq C$ (see definition 4.4). Then $(M,\boldsymbol{a}) \equiv_{\infty q} (M, f(\boldsymbol{a}))$ (see definition 4.2 for $\equiv_{\infty q}$ and §2.1 for $f(\boldsymbol{a})$).*

Proof. Write $f(\boldsymbol{a}) = \boldsymbol{b} = (b_i : i \in I)$. By lemma 4.3, it suffices to show that $\exists$ has a winning strategy in $\mathrm{EF}_q(M,\boldsymbol{a},M,\boldsymbol{b})$. Such a strategy can be outlined by saying that if $\forall$ plays an element 'close to' $\boldsymbol{a}$ or $\boldsymbol{b}$ then $\exists$ responds using f, but otherwise she just copies $\forall$'s move. The notion of 'close' is dynamic and depends on the round number and the earlier moves in the game. In a little more detail, adapting the summary of a similar game in [13, p.283]: $\exists$ merely needs to respect, in round t of the game, the critical distance $d_t = 2^{q-t}$; if $\forall$'s move in round t goes to within distance d_t (as measured by $d^{M,f}$) of either an element in $\boldsymbol{a}$ or $\boldsymbol{b}$ or an already-played element to which her response used f, then $\exists$ plays according to f; otherwise, $\exists$ responds by copying $\forall$'s move.

The details are actually very simple. For each ordinal $t < q$, we will write a_t and b_t for the 'left' and 'right' elements (respectively) chosen by the players in round t of this game. See §4.3 for the nomenclature. We assume without loss of generality that $I \cap q = \emptyset$. $\exists$ will play in round t as follows, and also ensure that the conditions R1_t–R4_t below hold at the start of round t, and that R1_{t+1}–R4_{t+1} hold at the end of round t (so also at the start of round $t+1$, when $t+1 < q$), where

$$d_t = 2^{q-t}.$$

Recall that $t = \{0, \ldots, t-1\}$ and $t + 1 = t \cup \{t\}$. We write d for $d^{M,f}$ in the proof.

R1_t Each element of $I \cup t$ is coloured either black or white.

R2_t Suppose that $i \in I \cup t$ is white. Then $b_i = a_i$.

R3_t Suppose that $j \in I \cup t$ is black. Then $a_j \in C$, $b_j = f(a_j)$, and $\{m \in M : d(m, a_j) < d_t\} \subseteq C$.

R4_t Suppose that $i, j \in I \cup t$ are white and black, respectively. Then $d(a_i, a_j) > d_t$.

These conditions are left–right symmetric in the following sense. For each $i \in I \cup t$ we have $a_i = b_i$ or $f(a_i) = b_i$, and either way, $d(a_i, b_i) = 0$ by definition of $d = d^{M,f}$ — definition 4.4 is formulated to get this. So by the triangle inequality, $d(m, a_i) = d(m, b_i)$ for every $m \in M$. It follows that R1_t–R4_t are equivalent to the versions in which each a_i is swapped with b_i, and f in R3_t is replaced by its inverse, f^{-1}.

For $t = 0$, $\exists$ colours each element of I black. Then R1_0 holds trivially, R2_0 and R4_0 vacuously, and R3_0 by the assumptions.

Let $t < q$ and assume inductively that R1_t–R4_t hold at the start of round t. Suppose that $\forall$ chooses a left element $a_t \in M$, say. (The argument when he chooses a right element b_t is similar because of the left–right symmetry of R1–R4; we need only swap all a_i with b_i and replace f by f^{-1} below.) $\exists$ must select a right element b_t in response, and establish R1_{t+1}–R4_{t+1}. There are two cases.

Case 1: $d(a_t, a_j) \leq d_{t+1} = d_t/2$ *for some black* $j \in I \cup t$. Then $\exists$ extends the colouring of $I \cup t$ given by R1_t to $I \cup (t+1)$, by colouring t black. Since $d(a_t, a_j) \leq d_{t+1} < d_t$, the last part of R$3_t$ gives $a_t \in C = \operatorname{dom} f$. $\exists$ responds to $\forall$'s move with $b_t = f(a_t)$.

We check that R1_{t+1}–R4_{t+1} hold. R1_{t+1} and R2_{t+1} are already clear. R3_{t+1} for black elements of $I \cup t$ follows from R3_t, since $d_{t+1} \leq d_t$. The new case is t. We already know that $a_t \in C$ and $b_t = f(a_t)$. For the last part, let $m \in M$ with $d(m, a_t) < d_{t+1}$. By the triangle inequality for d and the case assumption, $d(m, a_j) \leq d(m, a_t) + d(a_t, a_j) < 2d_{t+1} = d_t$, so by R$3_t$ we obtain $m \in C$ as required.

For R4_{t+1}, let $i \in I \cup t$ be white. We show that $d(a_i, a_t) > d_{t+1}$. If not, then as above, $d(a_i, a_j) \leq d(a_i, a_t) + d(a_t, a_j) \leq 2d_{t+1} = d_t$, contradicting R$4_t$. All other instances of R4_{t+1} follow from R4_t, since $d_t \geq d_{t+1}$.

Case 2: *otherwise.* This time, $\exists$ colours t white and sets $b_t = a_t$. So R1_{t+1} and R2_{t+1} obviously hold, and R3_{t+1} follows from R3_t as there are no new cases. The only new case to check in R4_{t+1} is that $d(a_t, a_j) > d_{t+1}$ whenever $j \in I \cup t$ is black — and this is exactly the case assumption.

$$\begin{array}{ccccccc} M & = & A & & + & & B \\ & & \ell \big\uparrow & & & & \big\uparrow r \\ & & A & \supseteq & E & \subseteq & B \end{array}$$

Figure 1: the L-structure M and embeddings ℓ, r in corollary 4.6

That completes the definition of $\exists$'s strategy. We check that it is winning. At the end of the game, R1_q–R4_q hold, and $d_q = 2^0 = 1$. Let $\alpha(x_1, \ldots, x_n)$ be an atomic L-formula (which may be an equality or involve constants), and $I' = \{i_1, \ldots, i_n\} \subseteq I \cup q$. We show that $M \models \alpha(a_{i_1}, \ldots, a_{i_n}) \leftrightarrow \alpha(b_{i_1}, \ldots, b_{i_n})$.

We can assume without loss of generality that $x_1, \ldots, x_n$ all occur in α. Suppose that $M \models \alpha(a_{i_1}, \ldots, a_{i_n})$. Then by definition of d,

$$d(a_i, a_j) \leq 1 = d_q \quad \text{for each } i, j \in I'.$$

It follows from R4_q that $i_1, \ldots, i_n$ all have the same colour. If this is white, then by R2_q, $b_i = a_i$ for each $i \in I'$, so obviously $M \models \alpha(b_{i_1}, \ldots, b_{i_n})$. If it is black, then $b_i = f(a_i)$ for each $i \in I'$, by R3_q. Since C is a substructure of M, the partial map $f : M \to M$ is a partial isomorphism, so again, $M \models \alpha(b_{i_1}, \ldots, b_{i_n})$.

The converse is similar, again using left–right symmetry of R1_q–R4_q. So $\exists$ won. □

Lest the proposition seem too abstract, here is a concrete instance of it that will be used later, in §5.2 and theorem 8.3.

COROLLARY 4.6 *Let L be a relational signature and A, B be L-structures having a common substructure E. Let $\boldsymbol{e} = (e_i : i \in I)$ be a family of elements of E, let $q < \omega$, and assume that*

$$\mathcal{N}^A_{2^q}(\boldsymbol{e}) \subseteq E \quad \textit{and} \quad \mathcal{N}^B_{2^q}(\boldsymbol{e}) \subseteq E. \tag{2}$$

Let $M = A + B$ (the disjoint union of A and B) and let $\ell : A \hookrightarrow M$ and $r : B \hookrightarrow M$ be the inclusion maps — see §4.1 and figure 1. Then $(M, \ell(\boldsymbol{e})) \equiv_{\infty q} (M, r(\boldsymbol{e}))$.

Proof (sketch). Let C be the substructure of M with domain $\ell(E) \cup r(E)$, and let f be the automorphism of C given by $f(\ell(e)) = r(e)$ and $f(r(e)) = \ell(e)$, for each $e \in E$. Let $\boldsymbol{a} = \ell(\boldsymbol{e})$ and $\boldsymbol{b} = r(\boldsymbol{e})$, so $f(\boldsymbol{a}) = \boldsymbol{b}$. The weighted graph $\mathfrak{G}(M, f)$ is $\mathfrak{G}(A) \uplus \mathfrak{G}(B)$ but with an edge of weight 0 added between $\ell(e)$ and $r(e)$ for each $e \in E$. In the light of this, the assumptions (2) easily yield $\mathcal{N}^{M,f}_{2^q}(\boldsymbol{a}) \subseteq C$. The conclusion $(M, \boldsymbol{a}) \equiv_{\infty q} (M, \boldsymbol{b})$ now follows by proposition 4.5. □

4.6 Interpretations

We will use *interpretations* in §5.3 to extend the reach of corollary 4.6. They 'interpret', 'define', or 'encode' a new structure in a given one. We broadly follow Hodges [14, §5.3] for the definitions. We will need infinitary interpretations with parameters, but only one-dimensional quantifier-free unrelativised ones. (Results for more general interpretations can also be obtained.) We pay attention to the syntactic side of interpretations because we want to use a single interpretation in multiple structures — see lemmas 4.7, 5.7, 5.8, and theorem 8.3.

Let K, L be signatures (which here have no function symbols, recall), let I be an index set, and let v_i $(i \in I)$ be new pairwise distinct variables taken not to occur in $L_{\infty\omega}$-formulas. An *interpretation of L in K with parameters I* (more fully, *with parameters v_i $(i \in I)$*) is a map $\mathcal{I}$ that provides a quantifier-free $K_{\infty\omega}$-formula $\mathcal{I}(\alpha)(x_1, \ldots, x_n, (v_i : i \in I))$ for each atomic L-formula $\alpha(x_1, \ldots, x_n)$. We extend $\mathcal{I}$ to all $L_{\infty\omega}$-formulas by induction in the obvious way: $\mathcal{I}(\neg\varphi) = \neg\mathcal{I}(\varphi)$, $\mathcal{I}(\bigwedge S) = \bigwedge\{\mathcal{I}(\varphi) : \varphi \in S\}$, similarly for $\bigvee S$, and $\mathcal{I}(\exists x\varphi) = \exists x\mathcal{I}(\varphi)$. Plainly, $\mathcal{I}(\varphi)$ always has the same quantifier depth as φ.

Now let A be a K-structure and $\boldsymbol{a} = (a_i : i \in I) \in {}^I A$. Let M be an L-structure with the same domain as A. We say that $\mathcal{I}$ *interprets M in $(A, \boldsymbol{a})$* if for each atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $m_1, \ldots, m_n \in M$, we have

$$M \models \alpha(m_1, \ldots, m_n) \quad \text{iff} \quad A \models \mathcal{I}(\alpha)(m_1, \ldots, m_n, (a_i : i \in I)).$$

Thus, for example, for each constant $c \in L$, by taking α above to be $x = c$, for each $m \in M$ we have $m = c^M$ iff $A \models \mathcal{I}(x = c)(m, \boldsymbol{a})$. There is clearly at most one M that $\mathcal{I}$ interprets in $(A, \boldsymbol{a})$, so when there is one, we can write it as $\mathcal{I}(A, \boldsymbol{a})$.

LEMMA 4.7 *Let $K, L, I, \mathcal{I}, A, \boldsymbol{a}$ be as above, and suppose that $\mathcal{I}(A, \boldsymbol{a})$ exists. Also suppose that B is a K-structure, $\boldsymbol{b} = (b_i : i \in I) \in {}^I B$, and $\mathcal{I}(B, \boldsymbol{b})$ exists.*

1. *For each $L_{\infty\omega}$-formula $\varphi(x)$ and $a \in A$,*

$$\mathcal{I}(A, \boldsymbol{a}) \models \varphi(a) \quad \textit{iff} \quad A \models \mathcal{I}(\varphi)(a, (a_i : i \in I)).$$

2. *If $f : A \to B$ is a K-embedding and $f(\boldsymbol{a}) = \boldsymbol{b}$, then $f : \mathcal{I}(A, \boldsymbol{a}) \to \mathcal{I}(B, \boldsymbol{b})$ is an L-embedding.*

3. *Let $j \in I$ and $q < \omega$. If $(A, \boldsymbol{a}) \equiv_{\infty q} (B, \boldsymbol{b})$, then $(\mathcal{I}(A, \boldsymbol{a}), a_j) \equiv_{\infty q} (\mathcal{I}(B, \boldsymbol{b}), b_j)$.*

Proof. (1) is straightforward by induction on φ, and follows from the 'reduction theorem' of [14, theorem 5.3.2]. For (2), assuming wlog. that $I \cap \{1, \ldots, n\} = \emptyset$, for each atomic L-formula $\alpha(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in A$ we have

	$\mathcal{I}(A, \boldsymbol{a}) \models \alpha(a_1, \ldots, a_n)$	
iff	$A \models \mathcal{I}(\alpha)(a_1, \ldots, a_n, (a_i : i \in I))$	by definition of $\mathcal{I}(A, \boldsymbol{a})$,
iff	$B \models \mathcal{I}(\alpha)(f(a_1), \ldots, f(a_n), (b_i : i \in I))$	since $\mathcal{I}(\alpha)$ is quantifier-free and f a K-embedding,
iff	$\mathcal{I}(B, \boldsymbol{b}) \models \alpha(f(a_1), \ldots, f(a_n))$	by definition of $\mathcal{I}(B, \boldsymbol{b})$.

For (3), let $\varphi(x)$ be any $L_{\infty\omega}$-formula of quantifier depth $\leq q$. Then the $K_{\infty\omega}$-formula $\mathcal{I}(\varphi)(x, (v_i : i \in I))$ also has quantifier depth $\leq q$. So

$\mathcal{I}(A, \boldsymbol{a}) \models \varphi(a_j)$	iff	$A \models \mathcal{I}(\varphi)(a_j, (a_i : i \in I))$	by part 1,
	iff	$B \models \mathcal{I}(\varphi)(b_j, (b_i : i \in I))$	since $(A, \boldsymbol{a}) \equiv_{\infty q} (B, \boldsymbol{b})$,
	iff	$\mathcal{I}(B, \boldsymbol{b}) \models \varphi(b_j)$	by part 1 for B. $\square$

$$
\begin{array}{ccccccccc}
(A,\mathsf{a}) & \sim & (A^l,\hat{\mathsf{a}}) & \sim & (\mathcal{I}(M,\boldsymbol{m}),\ell(\hat{\mathsf{a}})) & \equiv_{\infty q} & (\mathcal{I}(M,\boldsymbol{n}),r(\hat{\mathsf{a}})) & \sim & (A^l_{<l},\hat{\mathsf{a}}) \\
\wr^{\star}_{m} & & & & & \Longrightarrow & & & \wr^{\star} \\
(B,\mathsf{b}) & \sim & (B^l,\hat{\mathsf{b}}) & \sim & (\mathcal{I}'(M',\boldsymbol{m}'),\ell'(\hat{\mathsf{b}})) & \equiv_{\infty q} & (\mathcal{I}'(M',\boldsymbol{n}'),r'(\hat{\mathsf{b}})) & \sim & (B^l_{<l},\hat{\mathsf{b}})
\end{array}
$$

Figure 2: guide to our route

5 Unravellings

The first aim of this paper is to prove that, classically and in the finite and for each $\star \in \{\Diamond, \Diamond@\}$, every $\star$-bisimulation-invariant first-order $L(\sigma)$-formula $\varphi(x)$ is equivalent to a hybrid $\mathcal{L}^\star(\sigma)$-formula. We do it by showing that any two pointed Kripke models (A,a) and (B,b) for σ that agree on $\mathcal{L}^\star(\sigma)$-formulas up to a certain modal depth m, to be specified in definition 5.11, also agree on φ. Our approach to this is outlined in figure 2.

The models (A,a) and (B,b) can be seen on the left of the figure. The other models shown will be introduced later in this section. Suppose that φ has quantifier depth q, say. Then all models on the top row of figure 2 agree on φ — either because they are $\Diamond@$-bisimilar (written $\sim$) or because they agree on infinitary formulas of quantifier depth q (written $\equiv_{\infty q}$). The same goes for the bottom row. Further, if $(A,\mathsf{a}) \sim^\star_m (B,\mathsf{b})$ on the left then $(A^l_{<l},\hat{\mathsf{a}}) \sim^\star (B^l_{<l},\hat{\mathsf{b}})$ on the right, so these latter models agree on φ as well. Chasing φ right round figure 2 from top left to bottom left now shows that (A,a) and (B,b) agree on φ.

This is the core of the proof of theorem 6.1 in the next section. If A and B are finite, then so are all models in figure 2, and so the proof goes through in the finite.

The work of this current section is to define the models and establish the statements in figure 2. For the entire section, fix a hybrid signature $\sigma = \textsc{Prop} \uplus \textsc{Nom}$; all Kripke models will be for this signature. We write L for $L(\sigma)$ here. We use a, b below (and above) simply to distinguish them from other as and bs.

5.1 Unravelling a Kripke model

The modal notion of 'unravelling' a Kripke model is well known: see, e.g., [9, §3.2]. Here, we introduce and study a modified unravelling that works in the presence of nominals. We give an example after the definition. Until §5.4, fix an arbitrary pointed Kripke model (A,a) and $q < \omega$, and let $l = 2^q$.

DEFINITION 5.1 We define an L-structure (or Kripke model for σ) A^l from A. It is our 'hybrid depth-l unravelling' of A, and is finite if A is finite. First, some preliminaries.

- Let $N = \{c^A : c \in \textsc{Nom}\} \subseteq A$.
- For $k < \omega$, a *path of length k* (in A) is a sequence $(a_0, a_1, \ldots, a_k) \in {}^{k+1}A$, where $a_0 \in A$, $a_1, \ldots, a_k \in A \setminus N$, and $A \models a_i R a_{i+1}$ for each $i < k$. Only the first element of a path can lie in N.
- For a path $t = (a_0, \ldots, a_k)$ and $a \in A \setminus N$ with $A \models a_k R a$, we write $t^\frown a$ for the path $(a_0, \ldots, a_k, a)$.
- For $a \in A$, we usually write $\hat{a}$ for the path (a) of length 0. This is more compact. For $S \subseteq A$ we put $\hat{S} = \{\hat{s} : s \in S\}$.

- For $k < \omega$ let $\mathrm{Path}_{\leq k}(A)$ (resp., $\mathrm{Path}_{<k}(A)$) be the set of paths of length $\leq k$ (resp., of length $< k$) in A.

We now define A^l. Its domain is $\mathrm{Path}_{\leq l}(A)$. We read the symbols in σ according to the last elements of paths: so we define $A^l \models P((a_0, \ldots, a_k))$ iff $A, a_k \models p$ for each $p \in$ PROP, and for each nominal $c \in$ NOM we put $c^{A^l} = (c^A)$, a path in $\mathrm{Path}_{\leq l}(A)$ of length 0.

For the accessibility relation, let $t = (a_0, \ldots, a_k) \in \mathrm{Path}_{\leq l}(A)$. Then we define:

- $A^l \models tRu$ for each $u \in \mathrm{Path}_{\leq l}(A)$ of the form $t^\frown a$, where $a \in A \setminus N$.
- $A^l \models tR\hat{n}$ for each $n \in N$ with $A \models a_k Rn$.
- If $k = l$, then $A^l \models tR\hat{a}$ for each $a \in A$ with $A \models a_l Ra$.

These are the only instances of R. Finally let $\lambda : A^l \to A$ be the 'projection' function that maps each path $(a_0, \ldots, a_k)$ to its last element a_k.

The upshot of the definition of R is that 'R-arrows' in A^l can come into $\hat{N}$ from anywhere, and into $\hat{A} \setminus \hat{N}$ from paths of length l only. But each path of length $k \geq 1$ has a unique R-predecessor, namely, its initial segment (prefix) of length $k - 1$.

EXAMPLE 5.2 Consider the Kripke model $M = (W, R^M, V)$ for a signature comprising a single nominal, c, where $W = \{0, 1, 2\}$, $R^M = \{(0, 1), (1, 0), (1, 2), (2, 1)\}$, and $V(c) = \{2\}$. See figure 3. In figures 3–5, arrows indicate R-relations and undirected lines indicate R-relations going both ways.

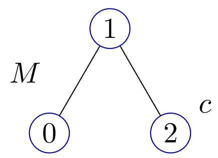

Figure 3: The model M

The modal depth-2 unravelling of $(M, 0)$, as per [19, definition 7], [17, §2.2], and [13, lemma 36], is shown in figure 4. It comprises all directed R-paths in M from 0 of length ≤ 2, with each path of length 2 identified with its endpoint in a new copy of M. Here and below, a path $(0, 1, 2)$, say, is written as just 012. There are no nominals in modal logic, so we treat c as an atom.

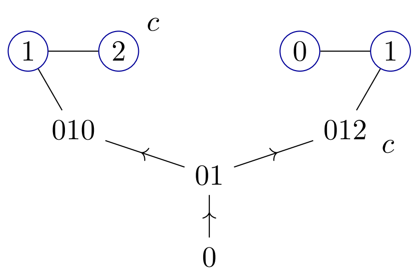

Figure 4: Classical modal unravelling to depth 2 of $(M, 0)$

This is plainly unsuitable for us: the 'nominal' c is true at two different points. This problem is fixed in the depth-2 unravelling M^2 in definition 5.1, as shown in figure 5 — c is true only at the length-0 path '2'.

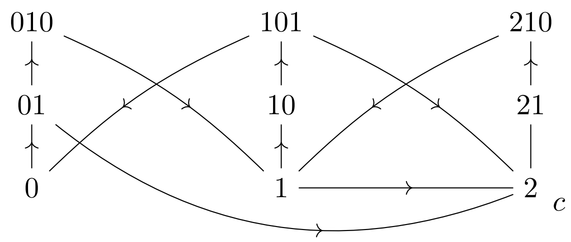

Figure 5: Our unravelling M^2 of M

Here are the key points of the construction in this case. Since $c^M = 2$, each path in figure 4 involving 2 as a non-initial point is 'pruned' by deleting all its points before the last 2. Here, the only such path is 012, which is pruned to 2. All R-arrows to and from paths are kept after pruning. The nominal c is made true just at the path 2. Now that we have 'backarrows' to this path, we may as well use length-0 paths instead of the copies of M in figure 4. This seems conceptually simpler. So we allow all pruned paths of length ≤ 2 starting not just at the point 0 of $(M, 0)$, but anywhere in M — this unravelling does not depend on the point. Each path ending in 1 has an R-arrow to the length-0 path 2, and each path yza of length 2 has an R-arrow to the length-0 path b for each $b \in M$ with $M \models aRb$.

The projection onto M of the model M^2 in figure 5, taking each path to its last point, is a bisimulation. We now show that this is true generally.

LEMMA 5.3 $(A, \mathsf{a}) \sim^{\Diamond @} (A^l, \hat{\mathsf{a}})$.

Proof. One can easily verify that $\lambda : A^l \to A$ is a $\Diamond @$-bisimulation and $\lambda(\hat{\mathsf{a}}) = \mathsf{a}$. We check only the 'Back' property. Suppose that $t = (a_0, \ldots, a_k) \in A^l$, so $k \leq l$ and $\lambda(t) = a_k$, and let $a \in A$ satisfy $A \models a_k R a$. We seek $u \in A^l$ with $A^l \models tRu$ and $\lambda(u) = a$. If $a \in N$ or $k = l$, take $u = \hat{a}$. If $a \in A \setminus N$ and $k < l$, take $u = t^\frown a$. $\square$

DEFINITION 5.4 We let $A^l_{<l}$ be the substructure of A^l with domain $\mathrm{Path}_{<l}(A)$. To reduce clutter, for λ as above, we write its restriction $\lambda \upharpoonright A^l_{<l}$ as $\lambda : A^l_{<l} \to A$ as well.

So $A^l_{<l}$ is obtained by simply deleting from A^l all paths of length l. The result is nonempty (since $l \geq 1$) and contains all elements of A^l named by constants (nominals), so is an L-structure. It can be 'well controlled' by hybrid logic, as lemma 5.12 will show. Perhaps we had better point out that $A^l_{<l}$ is not A^{l-1}, since the two give different meanings to R on paths of length $l-1$.

The restriction $\lambda : A^l_{<l} \to A$ is an L-homomorphism (see §4.1) and preserves atoms and nominals both ways, but it is not in general a bisimulation, because paths of length l have been deleted, so the Back property may fail.

5.2 Invoking locality

DEFINITION 5.5 We introduce a new relational signature K, obtained from L by deleting each constant c ($c \in \textsc{Nom}$) and adding a new unary relation symbol P_a for each $a \in A$. So K comprises R, a unary relation symbol P for each $p \in \textsc{Prop}$, and the new symbols P_a. It depends on A and may be infinite.

We define a K-structure $A^{l:K}$ with the same domain and interpretations of atoms in $\textsc{Prop}$ as A^l, and with

- $A^{l:K} \models tRu$ iff $A^l \models tRu$ and $u \notin \hat{A}$. (Hence, u is a path extending t by one.)
- $A^{l:K} \models P_a(t)$ iff $A^l \models tR\hat{a}$, for each $a \in A$. (Hence, $a \in N$ or t is a path of length l.)

We let $A^{l:K}_{<l}$ be the substructure of $A^{l:K}$ with domain $\mathrm{Path}_{<l}(A)$.

In $A^{l:K}$, we have removed all R-arrows into $\hat{A}$, but the P_a ensure that they are not forgotten. We also removed the nominals: their values will be remembered 'by hand'.

We are now going to use A as an index set. Write $\boldsymbol{a} = (\hat{a} : a \in A)$. This is a family of elements of each of A^l, $A^l_{<l}$, $A^{l:K}$, and $A^{l:K}_{<l}$. It is clear from the definitions (see definition 4.4(3) for $\mathcal{N}$) that[4]

$$\mathcal{N}_l^{A^{l:K}}(\boldsymbol{a}) = \mathcal{N}_l^{A^{l:K}_{<l}}(\boldsymbol{a}) = A^{l:K}_{<l}. \tag{3}$$

Let $M = A^{l:K} + A^{l:K}_{<l}$ (disjoint union), and let $\ell : A^{l:K} \hookrightarrow M$ and $r : A^{l:K}_{<l} \hookrightarrow M$ be the respective K-embeddings, as in corollary 4.6. Note that if A is finite then so is M. Write $\boldsymbol{m} = \ell(\boldsymbol{a})$ and $\boldsymbol{n} = r(\boldsymbol{a})$. By corollary 4.6 with $A = A^{l:K}$, $B = E = A^{l:K}_{<l}$, and $\boldsymbol{e} = \boldsymbol{a}$, we obtain

$$(M, \boldsymbol{m}) \equiv_{\infty q} (M, \boldsymbol{n}). \tag{4}$$

The corollary applies since K is relational and by (3). We are using only the special case of it where $B = E$. The general case will be used in theorem 8.3.

5.3 Invoking interpretations

To get back to L, we use an interpretation.

DEFINITION 5.6 We define an interpretation $\mathcal{I}$ of L in K with parameters A. It takes each atomic L-formula $\alpha(x_1, \ldots, x_n)$ to a quantifier-free $K_{\infty\omega}$-formula $\mathcal{I}(\alpha)(x_1, \ldots, x_n, (v_a : a \in A))$. Here recall §4.6 — the index set I there is A here, and the pairwise distinct variables v_a $(a \in A)$ are taken not to occur in $L_{\infty\omega}$-formulas.

1. Let $\mathcal{I}(x = y)$ be $x = y$, and let $\mathcal{I}(P(x))$ be $P(x)$ for $p \in$ PROP.
2. Let $\mathcal{I}(xRy)$ be $xRy \vee \bigvee_{a \in A}(P_a(x) \wedge y = v_a)$.
3. For an atomic L-formula $\alpha(x_1, \ldots, x_n, y_1, \ldots, y_m)$ not involving any constants, and constants (nominals) $c_1, \ldots, c_m \in L$, define $\mathcal{I}(\alpha(x_1, \ldots, x_n, c_1, \ldots, c_m))$ to be the result of substituting $v_{c_i^A}$ for y_i in the formula $\mathcal{I}(\alpha(x_1, \ldots, x_n, y_1, \ldots, y_m))$ defined above, for each $i = 1, \ldots, m$.

As an example, if $c \in$ NOM and $c^A = s \in A$, say, then $\mathcal{I}(x = c)$ is $x = v_s$; if $p \in$ PROP then $\mathcal{I}(P(c))$ is $P(v_s)$; and $\mathcal{I}(xRc)$ is $xRv_s \vee \bigvee_{a \in A}(P_a(x) \wedge v_s = v_a)$.

It should be clear that the L-structures $\mathcal{I}(A^{l:K}, \boldsymbol{a})$ and $\mathcal{I}(A^{l:K}_{<l}, \boldsymbol{a})$ exist and are A^l and $A^l_{<l}$, respectively. The L-structures $\mathcal{I}(M, \boldsymbol{m})$ and $\mathcal{I}(M, \boldsymbol{n})$ also exist. In contrast to M, they are not disjoint unions, because L is not relational (if NOM $\neq \emptyset$), and there may be 'R-arrows' running between $\ell(A^{l:K})$ and $r(A^{l:K}_{<l})$. Nonetheless, we have the following:

LEMMA 5.7 *$(A^l, \hat{\mathsf{a}}) \sim^{\Diamond @} (\mathcal{I}(M, \boldsymbol{m}), \ell(\hat{\mathsf{a}}))$ and $(A^l_{<l}, \hat{\mathsf{a}}) \sim^{\Diamond @} (\mathcal{I}(M, \boldsymbol{n}), r(\hat{\mathsf{a}}))$.*

[4] Without K, (3) can fail. For example, in M^2 as in figure 5, there is an R-arrow to 0 from 101. So the open radius-2 Gaifman neighbourhood of 0 contains 101 and so is not contained in $M^2_{<2}$. This is the main reason why we introduce $A^{l:K}$.

Proof. We show that $\ell : A^l \hookrightarrow \mathcal{I}(M, \boldsymbol{m})$ is a $\Diamond$@-bisimulation. We know that $\ell : A^{l:K} \hookrightarrow M$ is a K-embedding. By lemma 4.7(2), ℓ is also an L-embedding from $\mathcal{I}(A^{l:K}, \boldsymbol{a}) = A^l$ into $\mathcal{I}(M, \boldsymbol{m})$. It therefore preserves σ both ways, satisfies Forth, and is defined on all points named by nominals.

For Back, let $t \in A^l$ and $u \in M$, and suppose that $\mathcal{I}(M, \boldsymbol{m}) \models \ell(t)Ru$. We seek $t' \in A^l$ with $A^l \models tRt'$ and $\ell(t') = u$. We have $M \models \mathcal{I}(xRy)(\ell(t), u, (m_a : a \in A))$ by definition of $\mathcal{I}(M, \boldsymbol{m})$. So by definition of $\mathcal{I}(xRy)$,

$$M \models \ell(t)Ru \vee \bigvee_{a \in A} (P_a(\ell(t)) \wedge u = m_a).$$

But each disjunct here implies $u \in \ell(A^l)$: the first by definition of M as a disjoint union, and the others since $m_a = \ell(\hat{a})$ for each $a \in A$. So let $t' = \ell^{-1}(u) \in A^l$. Then $\mathcal{I}(M, \boldsymbol{m}) \models \ell(t)R\ell(t')$, and as ℓ is an L-embedding, $A^l \models tRt'$, as required. Essentially we proved that $\ell(A^l)$ is a generated submodel of $\mathcal{I}(M, \boldsymbol{m})$.

Similarly we can show that $r : A^l_{<l} \hookrightarrow \mathcal{I}(M, \boldsymbol{n})$ is a $\Diamond$@-bisimulation. □

LEMMA 5.8 $(\mathcal{I}(M, \boldsymbol{m}), \ell(\hat{\mathsf{a}})) \equiv_{\infty q} (\mathcal{I}(M, \boldsymbol{n}), r(\hat{\mathsf{a}}))$.

Proof. By (4) and lemma 4.7(3), taking 'j' there to be a here, so the jth entry of the family $\boldsymbol{m} = (\ell(\hat{a}) : a \in A)$ is $\ell(\hat{\mathsf{a}})$, and the jth entry of $\boldsymbol{n} = (r(\hat{a}) : a \in A)$ is $r(\hat{\mathsf{a}})$. □

5.4 Summary so far

We have proved the following:

PROPOSITION 5.9 *Let (A, a) be a pointed Kripke model, $q < \omega$, and $l = 2^q$. Then:*

$$(A, \mathsf{a}) \sim^{\Diamond @} (A^l, \hat{\mathsf{a}}) \sim^{\Diamond @} (\mathcal{I}(M, \boldsymbol{m}), \ell(\hat{\mathsf{a}})) \equiv_{\infty q} (\mathcal{I}(M, \boldsymbol{n}), r(\hat{\mathsf{a}})) \sim^{\Diamond @} (A^l_{<l}, \hat{\mathsf{a}}).$$

If A is finite then so are all the structures here.

Proof. By lemmas 5.3, 5.7, 5.8, and 5.7, respectively. Finiteness has already been discussed. □

Applying the proposition to (A, a) and (B, b) establishes the top and bottom lines of figure 2. This will be shown formally in theorem 6.1. We used $\mathcal{I}', M', \ldots$ in the bottom line of the figure because the items are defined using B and so are distinct from those on the top line.

5.5 Two connecting lemmas

So far, we have looked at unravellings of a single pointed Kripke model (A, a). Our final two lemmas draw out connections between the unravellings of two pointed Kripke models. They will establish the rest of figure 2. The reader may wish to review the bisimulation game $\mathrm{Bs}^\star_\alpha(A, \mathsf{a}, B, \mathsf{b})$ defined in §3.2. By definition 3.4, $(A, \mathsf{a}) \sim^\star_\alpha (B, \mathsf{b})$ means that $\exists$ has a winning strategy in this game.

LEMMA 5.10 *Let $(A, \mathsf{a}), (B, \mathsf{b})$ be pointed Kripke models, $\star \in \{\Diamond, \Diamond @\}$, $\alpha \leq \omega$, and $1 \leq l < \omega$. Suppose that $(A, \mathsf{a}) \sim^\star_\alpha (B, \mathsf{b})$. Then $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star_\alpha (B^l_{<l}, \hat{\mathsf{b}})$.*

Proof. Recall (from just after definition 5.4) that the projection $\lambda : A^l_{<l} \to A$ is an L-homomorphism (so preserves R forwards), and preserves atoms and nominals both ways. There is a similar projection $\mu : B^l_{<l} \to B$ taking each path in $B^l_{<l}$ to its last element. They may not be bisimulations.

Assume that $\exists$ has a winning strategy in $\mathrm{Bs}^\star_\alpha(A, \mathsf{a}, B, \mathsf{b})$. She will play this game privately, using her winning strategy, to help her win the main game $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$.

In a play of this main game, let the successive positions be $(t_0, u_0), \ldots, (t_s, u_s), \ldots$, say. $\exists$ will ensure that

(i) $t_s \in A^l_{<l}$ and $u_s \in B^l_{<l}$ are paths of equal length, for each s,

(ii) $(\lambda(t_0), \mu(u_0)), \ldots, (\lambda(t_s), \mu(u_s)), \ldots$ are successive positions in a play of the private game $\mathrm{Bs}^\star_\alpha(A, \mathsf{a}, B, \mathsf{b})$ in which she is using her winning strategy.

If she can do this, then since her strategy is winning, $(A, \lambda(t_s))$ and $(B, \mu(u_s))$ agree on σ, for each s. So by our opening remarks, $(A^l_{<l}, t_s)$ and $(B^l_{<l}, u_s)$ also agree on σ for each s, and $\exists$ will win.

We now explain how she can do it.

Suppose that $\forall$ chooses (t_0, u_0) as the initial position in the main game $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. If $(t_0, u_0) = (\hat{\mathsf{a}}, \hat{\mathsf{b}})$, then clearly, t_0 and u_0 have equal length 0, $\lambda(t_0) = \mathsf{a}$, and $\mu(u_0) = \mathsf{b}$. If $\star = \Diamond @$ and $(t_0, u_0) = (c^{A^l_{<l}}, c^{B^l_{<l}}) = ((c^A), (c^B))$ for some $c \in \textsc{Nom}$, then again t_0, u_0 have length 0, and $\lambda(t_0) = c^A$ and $\mu(u_0) = c^B$. So in all cases it is legal for $\forall$ to choose $(\lambda(t_0), \mu(u_0))$ for the initial position in the private game $\mathrm{Bs}^\star_\alpha(A, \mathsf{a}, B, \mathsf{b})$. $\exists$ lets him do so — she makes this choice on his behalf. Then conditions (i) and (ii) above are met.

In round $s < \alpha$ of the main game $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$, assume that $\exists$ has kept the two conditions so far, and suppose that $\forall$ plays $t_{s+1} \in A^l_{<l}$ (the argument is similar if he plays in $B^l_{<l}$; and if he cannot move then $\exists$ wins at this point and we are done). By the game rules, $A^l_{<l} \models t_s R t_{s+1}$.

Write $a = \lambda(t_{s+1})$. As λ is a homomorphism, $A \models \lambda(t_s) R a$ too, and it is legal for $\forall$ to play a in round s of the private game $\mathrm{Bs}^\star_\alpha(A, \mathsf{a}, B, \mathsf{b})$. Again, $\exists$ lets him do it, and responds using her winning strategy with $b \in B$, say. As her strategy is winning, (†) $B \models \mu(u_s) R b$ and (‡) (A, a) and (B, b) agree on σ.

We now define $\exists$'s response $u_{s+1} \in B^l_{<l}$ to $\forall$ in the main game $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. There are two cases. Suppose first that t_{s+1} is named in $A^l_{<l}$ by some nominal c. Then $t_{s+1} = (c^A)$, a path of length 0. Plainly, $a = c^A$. By (‡), $b = c^B$. Then $\exists$ lets $u_{s+1} = \hat{b} \in B^l_{<l}$, also a path of length 0. The reader can check that this is well defined if t_{s+1} is named by more than one nominal.

Now suppose otherwise, and let the common length of t_s, u_s be n, say. Then $t_{s+1} = t_s{}^\frown a$ by definition of $A^l_{<l}$. This path has length $n+1$, and $n+1 < l$ because $t_{s+1} \in A^l_{<l}$. Because λ preserves nominals both ways, a is not named by a nominal. By (‡), neither is b, so $u_s{}^\frown b$ is a path in B. It has length $n+1$ as well, so is in $B^l_{<l}$ since $n+1 < l$. $\exists$ lets $u_{s+1} = u_s{}^\frown b$.

In each case, $\exists$ has found $u_{s+1} \in B^l_{<l}$ of the same path length (0 or $n+1$) as t_{s+1}, with $\mu(u_{s+1}) = b$, and with $B^l_{<l} \models u_s R u_{s+1}$ (by (†) and the definition of $B^l_{<l}$). So $\exists$ can legally respond to $\forall$ with u_{s+1} in the main game $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$, and in doing so, keep conditions (i) and (ii) above. Hence, we have described a winning strategy for her in $\mathrm{Bs}^\star_\alpha(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. □

DEFINITION 5.11 For $\star \in \{\Diamond, \Diamond @\}$, $1 \leq l < \omega$, and $n < \omega$, define

$$\mathsf{m}(\star, l, n) = \begin{cases} l - 1, & \text{if } n = 0, \\ l, & \text{if } n > 0 \text{ and } \star = \Diamond @, \\ l \cdot (n+1), & \text{if } n > 0 \text{ and } \star = \Diamond. \end{cases}$$

LEMMA 5.12 *Suppose that* NOM *is finite, let* $\star \in \{\Diamond, \Diamond@\}$ *and* $1 \leq l < \omega$, *and write* $m = \mathsf{m}(\star, l, |\text{NOM}|)$. *Let* $(A, \mathsf{a}), (B, \mathsf{b})$ *be pointed Kripke models satisfying* $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star_m (B^l_{<l}, \hat{\mathsf{b}})$. *Then* $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star (B^l_{<l}, \hat{\mathsf{b}})$.

Proof. By an *R-chain of length $n < \omega$ in $A^l_{<l}$*, we will mean a sequence $t_0, \ldots, t_n$ of elements of $A^l_{<l}$ with $A^l_{<l} \models t_i R t_{i+1}$ for each $i < n$. (The word 'path' could be confusing here.) By construction of $A^l_{<l}$, if $n \geq l$ then at least one of $t_1, \ldots, t_n$ is named by a nominal (t_0 may be as well). An *R-chain in $B^l_{<l}$* is defined similarly.

By lemma 3.5(4), it is enough to show that $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star_\omega (B^l_{<l}, \hat{\mathsf{b}})$. By assumption, $\exists$ has a winning strategy in $\mathrm{Bs}^\star_m(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. She will use this strategy in a private play of this game, to help her win $\mathrm{Bs}^\star_\omega(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. Of course, it may run out, so she may have to reset it frequently.

There are three cases, according to how $m = \mathsf{m}(\star, l, |\text{NOM}|)$ is defined.

First take the case NOM $= \emptyset$, so $m = l - 1$. Then each play of $\mathrm{Bs}^\star_\omega(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$ comes to an end after at most m rounds, since no R-chain in $A^l_{<l}$ or $B^l_{<l}$ is longer than this when NOM $= \emptyset$. Hence, $\exists$ can just use her winning strategy in $\mathrm{Bs}^\star_m(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. (Since NOM $= \emptyset$, the games do not depend on $\star$.)

Now take the case when NOM $\neq \emptyset$ and $\star = \Diamond@$, so $m = l$. By assumption and lemma 3.5(1), $\exists$ has winning strategies in $\mathrm{Bs}^\Diamond_l(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$ and $\mathrm{Bs}^\Diamond_l(A^l_{<l}, (c^A), B^l_{<l}, (c^B))$ for each $c \in$ NOM. She can use them repeatedly in a play of $\mathrm{Bs}^{\Diamond@}_\omega(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$, as follows. Initially, she chooses whichever strategy matches $\forall$'s choice of initial position. As play of $\mathrm{Bs}^{\Diamond@}_\omega$ continues, consideration of the form of $A^l_{<l}$ and $B^l_{<l}$ shows that it will either end with a win for $\exists$ because $\forall$ can't move, or will arrive after $\leq l$ rounds at a position of the form $((c^A), (c^B))$ for some nominal c. $\exists$ can then pick up a winning strategy in $\mathrm{Bs}^\Diamond_l(A^l_{<l}, (c^A), B^l_{<l}, (c^B))$ lasting another l rounds. Continuing in this way, she will win. (Actually this argument works for infinite NOM so long as we still define $\mathsf{m}(\Diamond@, l, |\text{NOM}|) = l$.)

Finally take the case when NOM $\neq \emptyset$ and $\star = \Diamond$, so $m = l \cdot (|\text{NOM}| + 1)$. Let us say that elements $t \in A^l_{<l}$ and $u \in B^l_{<l}$ *match* if there is a finite R-chain in $A^l_{<l}$ running from $\hat{\mathsf{a}}$ to t, and for some nominal c we have $t = (c^A)$ and $u = (c^B)$.

Claim. If t, u match, then $\exists$ has a winning strategy in $\mathrm{Bs}^\Diamond_l(A^l_{<l}, t, B^l_{<l}, u)$.
Proof of claim. Take a shortest possible R-chain $t_0, \ldots, t_n$ in $A^l_{<l}$ from $\hat{\mathsf{a}}$ to t, so $t_0 = \hat{\mathsf{a}}$ and $t_n = t$. By minimality, $t_0, \ldots, t_n$ are pairwise distinct, and at most $|\text{NOM}|$ of them are named by a nominal. But whenever $0 \leq s < s + l \leq n$, some point in $\{t_{s+1}, \ldots, t_{s+l}\}$ is named by a nominal. It follows that $n \leq l \cdot |\text{NOM}|$.

By following the chain, playing in $A^l_{<l}$ in each round, $\forall$ can get from $\hat{\mathsf{a}}$ to t in $\leq l \cdot |\text{NOM}|$ rounds of a play of $\mathrm{Bs}^\Diamond_m(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. If $\exists$ uses her winning strategy in such a play, then as it preserves nominals, she will for sure arrive at u, and her winning strategy will still have $\geq m - l \cdot |\text{NOM}| = l$ rounds left to run. So 'continue with the strategy in progress' is a winning strategy for her in $\mathrm{Bs}^\Diamond_l(A^l_{<l}, t, B^l_{<l}, u)$. This proves the claim.

We finish as in the preceding case. Let $\forall, \exists$ play $\mathrm{Bs}^\Diamond_\omega(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$, with $\exists$ initially using her winning strategy in $\mathrm{Bs}^\Diamond_m(A^l_{<l}, \hat{\mathsf{a}}, B^l_{<l}, \hat{\mathsf{b}})$. Play will either end with a win for $\exists$ because $\forall$ can't move, or arrive after $\leq l$ rounds at a position (t, u) of the form $((c^A), (c^B))$ for some nominal c. Plainly, t and u will then match. So by the claim, $\exists$ can then pick up a winning strategy lasting another l rounds. Continuing in this way forever, she will win. $\square$

6 Characterisation theorem

THEOREM 6.1 *Let σ be a hybrid signature, let $\mathcal{C}$ be either the class of all pointed Kripke models for σ, or the class of all finite pointed Kripke models for σ, and let $\star \in \{\Diamond, \Diamond @\}$. Let $\varphi(x)$ be a first-order $L(\sigma)$-formula of quantifier depth q and involving n distinct constants (nominals) from $L(\sigma)$. Assume that φ is $\star$-bisimulation invariant over $\mathcal{C}$. Then φ is equivalent over $\mathcal{C}$ to some $\mathcal{L}^\star(\sigma)$-formula ψ of modal depth at most $m = \mathsf{m}(\star, 2^q, n)$ (see definition 5.11).*

Proof. Write $\sigma = \text{PROP} \cup \text{NOM}$. Let $\tau \subseteq \sigma$ be the $\subseteq$-least hybrid signature such that φ is an $L(\tau)$-formula. It simply collects all symbols of σ that actually occur in φ (after changing $p \in \text{PROP}$ to $P \in L(\sigma)$). First we prove the theorem assuming that $\sigma = \tau$. So σ is finite and $|\text{NOM}| = n$.

Claim 1. If $(A, \mathsf{a}), (B, \mathsf{b}) \in \mathcal{C}$ and $(A, \mathsf{a}) \sim^\star_m (B, \mathsf{b})$ then $A \models \varphi(\mathsf{a})$ iff $B \models \varphi(\mathsf{b})$.
Proof of claim. Write $l = 2^q$. As φ has quantifier depth q and is $\star$-bisimulation invariant over $\mathcal{C}$, by proposition 5.9 we have

$$A \models \varphi(\mathsf{a}) \iff A^l_{<l} \models \varphi(\hat{\mathsf{a}}). \tag{5}$$

Note here that since $(A, \mathsf{a}) \in \mathcal{C}$, all structures mentioned in the proposition are also in $\mathcal{C}$, and so each of its four steps preserves φ. Similarly, applying proposition 5.9 to (B, b) gives

$$B \models \varphi(\mathsf{b}) \iff B^l_{<l} \models \varphi(\hat{\mathsf{b}}). \tag{6}$$

By assumption, $(A, \mathsf{a}) \sim^\star_m (B, \mathsf{b})$, so by lemma 5.10, $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star_m (B^l_{<l}, \hat{\mathsf{b}})$, and thus $(A^l_{<l}, \hat{\mathsf{a}}) \sim^\star (B^l_{<l}, \hat{\mathsf{b}})$ by lemma 5.12 and since $m = \mathsf{m}(\star, l, |\text{NOM}|)$. As already observed, these structures are in $\mathcal{C}$. As φ is $\star$-bisimulation invariant over $\mathcal{C}$, we obtain

$$A^l_{<l} \models \varphi(\hat{\mathsf{a}}) \iff B^l_{<l} \models \varphi(\hat{\mathsf{b}}). \tag{7}$$

Putting (5)–(7) together proves the claim.

The rest of the proof is quite standard. Since σ is finite, we can form the finite set $\mathcal{F}^\star_m$ of $\mathcal{L}^\star(\sigma)$-formulas from definition 3.6. They have modal depth $\leq m$. Define

$$\begin{aligned} \mathrm{tp}(A, \mathsf{a}) &= \{\psi \in \mathcal{F}^\star_m : A, \mathsf{a} \models \psi\} \cup \{\neg\psi : \psi \in \mathcal{F}^\star_m,\ A, \mathsf{a} \models \neg\psi\}, \quad \text{for } (A, \mathsf{a}) \in \mathcal{C}, \\ \psi &= \bigvee \big\{ \bigwedge \mathrm{tp}(B, \mathsf{b}) : (B, \mathsf{b}) \in \mathcal{C},\ B \models \varphi(\mathsf{b}) \big\}. \end{aligned}$$

Although $\mathcal{C}$ is a proper class, the class following the disjunction is finite since $\mathcal{F}^\star_m$ is finite, so ψ is an $\mathcal{L}^\star(\sigma)$-formula of modal depth $\leq m$.

Claim 2. φ is equivalent over $\mathcal{C}$ to ψ.
Proof of claim. Let $(A, \mathsf{a}) \in \mathcal{C}$. If $A \models \varphi(\mathsf{a})$ then $\bigwedge \mathrm{tp}(A, \mathsf{a})$ is a disjunct of ψ, and plainly $A, \mathsf{a} \models \bigwedge \mathrm{tp}(A, \mathsf{a})$, so $A, \mathsf{a} \models \psi$.

Conversely, assume that $A, \mathsf{a} \models \psi$. So there is some $(B, \mathsf{b}) \in \mathcal{C}$ with $B \models \varphi(\mathsf{b})$ and $A, \mathsf{a} \models \bigwedge \mathrm{tp}(B, \mathsf{b})$. It follows by definition of $\mathrm{tp}(B, \mathsf{b})$ that (A, a) and (B, b) agree on $\mathcal{F}^\star_m$. By lemma 3.7,

$$(A, \mathsf{a}) \sim^\star_m (B, \mathsf{b}). \tag{8}$$

Since $B \models \varphi(\mathsf{b})$, claim 1 and (8) yield $A \models \varphi(\mathsf{a})$, proving the claim, and the theorem when $\sigma = \tau$.

Now we prove the theorem without restrictions on σ. Let $\mathcal{C}_\tau$ be the class of all (finite, if the models in $\mathcal{C}$ are finite) pointed Kripke models for τ. As φ is assumed $\star$-bisimulation invariant over $\mathcal{C}$, by lemma 4.1 it is also $\star$-bisimulation invariant over $\mathcal{C}_\tau$. So by the case of the theorem already proved, φ is equivalent over $\mathcal{C}_\tau$ to an $\mathcal{L}^\star(\tau)$-formula ψ of modal depth $\leq m$. But of course, φ is equivalent to ψ over $\mathcal{C}$ as well. For let $(A, \mathsf{a}) \in \mathcal{C}$. Because they agree on symbols in φ, we have $A \models \varphi(\mathsf{a})$ iff $A \upharpoonright \tau \models \varphi(\mathsf{a})$. As $A \upharpoonright \tau \in \mathcal{C}_\tau$, this is iff $A \upharpoonright \tau, \mathsf{a} \models \psi$. Because $A \upharpoonright \tau$ and A agree on symbols in ψ, this is iff $A, \mathsf{a} \models \psi$, as required. □

The case $\text{NOM} = \emptyset$ is just the modal case, and is well known, as made clear in the introduction. We include it to indicate how, and (see lemma 5.12) why, the bound on the modal depth of the equivalent formula varies with the choice of language.

7 Optimality of modal depth bounds

Theorem 6.1 showed that every first-order $L(\sigma)$-formula $\varphi(x)$ of quantifier depth q, written with n nominals, and $\star$-bisimulation-invariant over $\mathcal{C}$, is equivalent over $\mathcal{C}$ to a $\mathcal{L}^\star(\sigma)$-formula ψ of modal depth $\leq \mathsf{m}(\star, 2^q, n)$. Perhaps surprisingly for a model-theoretic method, but less so in the light of Otto's work, this bound is optimal. Of course, sometimes one can find a simpler ψ, but in the worst case one cannot. We now give examples to show this. Take $\sigma = \text{PROP} \uplus \text{NOM}$ with $\text{PROP} = \{p\}$. It makes no difference which $\mathcal{C}$ in the theorem is chosen.

First consider the case $\text{NOM} = \emptyset$, when of course $\mathcal{L}^\Diamond(\sigma) = \mathcal{L}^{\Diamond @}(\sigma)$ is the ordinary modal language. This case was dealt with by Otto, who mentioned in [17, exercise 3.1] (also with Goranko in [13, p.283]), and showed in elegant detail in [18, corollary 3.6], that for each $q < \omega$, the modal formula $\psi = \bigvee_{i<2^q} \Diamond^i p$ (see §2.2 for the definition) is equivalent over $\mathcal{C}$ to a first-order $L(\sigma)$-formula $\varphi(x)$ of quantifier depth q. Clearly, ψ has modal depth $2^q - 1 = \mathsf{m}(\Diamond, 2^q, 0) = \mathsf{m}(\Diamond @, 2^q, 0)$. To paraphrase [18], $\varphi(x)$ is not invariant under $\sim^\Diamond_\ell$ for any $\ell < 2^q - 1$, hence not equivalent over $\mathcal{C}$ to any modal formula of depth less than $2^q - 1$.

To help with the other cases, for $q < \omega$ define a first-order $L(\sigma)$-formula '$xR^{2^q}y$' of quantifier depth q by induction: $xR^{2^0}y$ is xRy, and $xR^{2^{q+1}}y$ is $\exists z(xR^{2^q}z \wedge zR^{2^q}y)$.

EXAMPLE 7.1 Let $\text{NOM} = \{c\}$ and $\star = \Diamond @$. Let $q < \omega$ and $l = 2^q$, so $\mathsf{m}(\star, l, 1) = l$. Define $\varphi(x) = cR^l c$, an $L(\sigma)$-formula of quantifier depth q. Over $\mathcal{C}$, φ is expressible in $\mathcal{L}^{\Diamond @}(\sigma)$ by $@_c \Diamond^l c$, of modal depth l.

To show that l is optimal, define finite Kripke models $A_1 = (\{a, 0, \ldots, l-1\}, R_1, V_1)$ and $A_2 = (\{a, 0, \ldots, l-1\}, R_2, V_1)$, where $a \notin \{0, \ldots, l-1\}$, $R_1 = \{(i, i+1) : i < l-1\}$, $R_2 = R_1 \cup \{(l-1, 0)\}$, and $V_1(c) = \{0\}$. Then $(A_1, a), (A_2, a) \in \mathcal{C}$. The strategy 'copy $\forall$'s moves' is winning for $\exists$ in $\mathrm{Bs}^{\Diamond @}_{l-1}(A_1, a, A_2, a)$, because the difference in the models is too far away from 0 to reach in $< l$ rounds. So by lemma 3.7, (A_1, a) and (A_2, a) agree on all $\mathcal{L}^{\Diamond @}(\sigma)$-formulas of modal depth $< l$. But clearly $A_1 \models \neg\varphi(a)$ and $A_2 \models \varphi(a)$, so $\varphi(x)$ is not equivalent over $\mathcal{C}$ to any $\mathcal{L}^{\Diamond @}(\sigma)$-formula of modal depth $< l$ — nor obviously to any $\mathcal{L}^\Diamond(\sigma)$-formula without @, since $(A_1, a) \sim^\Diamond (A_2, a)$, so $\varphi(x)$ is not $\Diamond$-bisimulation invariant.

EXAMPLE 7.2 Finally let $\text{NOM} = \{c_1, \ldots, c_n\}$, where $c_1, \ldots, c_n$ are pairwise distinct, and $\star = \Diamond$. Let $q < \omega$ and $l = 2^q$, so $\mathsf{m}(\star, l, n) = l(n+1) = m$, say. Define

$$\varphi(x) = xR^l c_1 \wedge c_1 R^l c_2 \wedge \ldots \wedge c_{n-1} R^l c_n \wedge c_n R^l c_1.$$

Again, φ has quantifier depth q. It is equivalent over $\mathcal{C}$ to the $\mathcal{L}^\Diamond(\sigma)$-formula

$$\Diamond^l (c_1 \wedge \Diamond^l (c_2 \wedge \cdots \wedge \Diamond^l (c_n \wedge \Diamond^l c_1)) \cdots).$$

This has modal depth $l(n+1) = m$.

Define finite Kripke models $A_3 = (\{0, \ldots, m-1\}, R_3, V_3)$ and $A_4 = (\{0, \ldots, m-1\}, R_4, V_3)$, where $R_3 = \{(i, i+1) : i < m-1\}$, $R_4 = R_3 \cup \{(m-1, l)\}$, and $V_3(c_i) = \{li\}$ for $i = 1, \ldots, n$. Then $(A_3, 0), (A_4, 0) \in \mathcal{C}$, $A_3 \models \neg\varphi(0)$, $A_4 \models \varphi(0)$, and again $\exists$ has the winning strategy 'copy $\forall$'s moves' in $\mathrm{Bs}^{\Diamond}_{m-1}(A_3, 0, A_4, 0)$, so lemma 3.7 yields that $(A_3, 0)$ and $(A_4, 0)$ agree on all $\mathcal{L}^{\Diamond}(\sigma)$-formulas of modal depth $< m$. Hence, $\varphi(x)$ is not equivalent over $\mathcal{C}$ to any such formula.

Some may have been surprised when we defined $d(@_c\psi) = d(\psi)$ (rather than $1 + d(\psi)$) in the definition of modal depth in §2.2. So whilst we gave a bound on the nesting depth of $\Diamond$s in the $\mathcal{L}^{\Diamond @}(\sigma)$-formula ψ equivalent to $\varphi(x)$, perhaps ψ has @s nested to a much greater depth? The answer is 'no'. We obtained ψ as a boolean combination of formulas in $\mathcal{F}^{\Diamond @}_m$ as in definition 3.6, and each formula in this set has at most one occurrence of @, so the '@-nesting depth' of ψ is at most 1.

8 Extensions

Theorem 6.1 applies to the class of all pointed Kripke models (for a given signature) and the class of all finite ones, for proto-hybrid logic and basic hybrid logic. We now consider briefly whether the theorem extends to some other classes and logics considered by Otto in [17, 18].

8.1 Elementary classes

Van Benthem's classical proof can be used to extend theorem 6.1 to any *elementary class* of pointed Kripke models for any σ, though the argument does not provide any modal depth bounds for ψ.

8.2 Bisimulation-closed classes

Following Otto [17], we now consider *bisimulation-closed classes.*

DEFINITION 8.1 For $\star \in \{\Diamond, \Diamond @\}$, a class $\mathcal{C}$ of pointed Kripke models (for some hybrid signature) is said to be *closed under $\star$-bisimulation* if $(A, a) \sim^\star (B, b) \in \mathcal{C}$ implies $(A, a) \in \mathcal{C}$.

For modal logic, Otto states the following in [17, corollary 4.1] (we paraphrase):

> Let $\mathcal{C}$ be a class of pointed Kripke models closed under bisimulation, and $\mathcal{C}_{\mathit{fin}}$ the class of finite structures within $\mathcal{C}$. Then a first-order formula $\varphi(x)$ of quantifier depth q is invariant under bisimulation over $\mathcal{C}$ [over $\mathcal{C}_{\mathit{fin}}$] iff $\varphi(x)$ is logically equivalent over $\mathcal{C}$ [over $\mathcal{C}_{\mathit{fin}}$] to a modal formula of modal depth $\leq 2^q - 1$.

We now show that this positive result fails for proto-hybrid logic, but does generalise to basic hybrid logic.

8.2.1 Proto-hybrid logic

Here we show that there is no bisimulation characterisation of proto-hybrid logic over $\Diamond$-bisimulation-closed classes, either classically or in the finite. Here, 'bisimulation' will mean '$\Diamond$-bisimulation'.

EXAMPLE 8.2 We exhibit a class $\mathcal{C}$ of pointed Kripke models that is closed under bisimulation, and a first-order formula $\varphi(x)$ that is bisimulation invariant over $\mathcal{C}$, but not equivalent even over its subclass $\mathcal{C}_{fin}$ of finite models to any $\mathcal{L}^{\Diamond}$-formula.

All Kripke models here will be for the hybrid signature σ with a single nominal, c. We briefly allow arbitrary (that is, possibly infinite) disjunctions of $\mathcal{L}^{\Diamond}(\sigma)$-formulas. They have the usual semantics, and they are plainly bisimulation invariant.

A pointed Kripke model (A, a) is said to be *farsighted* if $A, a \models \bigvee_{n<\omega} \Diamond^n c$. This formula is bisimulation invariant, so the class $\mathcal{C}$ of farsighted pointed Kripke models is closed under bisimulation.

Now let $\varphi(x)$ be the $L(\sigma)$-formula cRc. Over farsighted models, φ is equivalent to $\bigvee_{n<\omega} \Diamond^n(c \wedge \Diamond c)$, so φ is bisimulation invariant over $\mathcal{C}$. (It is obviously not bisimulation invariant in general.)

Suppose for contradiction that $\varphi(x)$ is equivalent to some $\mathcal{L}^{\Diamond}(\sigma)$-formula ψ over the class $\mathcal{C}_{fin}$ of finite models within $\mathcal{C}$. Let $n = d(\psi)$, and define the following finite models:

$$\left.\begin{array}{rcl} A_n &=& (\{0, \ldots, n\}, \{(i, i+1) : i < n\} \cup \{(n, n)\}, V_n) \\ B_n &=& (\{0, \ldots, n\}, \{(i, i+1) : i < n\}, V_n) \end{array}\right\} \quad \text{where } V_n(c) = \{n\}.$$

Then $(A_n, 0) \sim_n^{\Diamond} (B_n, 0)$, so by lemma 3.7, $(A_n, 0)$ and $(B_n, 0)$ agree on ψ. But $A_n \models \varphi(0)$ and $B_n \models \neg\varphi(0)$. Since $(A_n, 0), (B_n, 0) \in \mathcal{C}_{fin}$, this is a contradiction.

So where does the proof of theorem 6.1 go wrong for $\mathcal{C}$ and $\mathcal{C}_{fin}$? The answer is of course that claim 1 in the proof fails. The quantifier depth q of φ above is zero. Let $l = 2^q = 1$ and $m = \mathsf{m}(\Diamond, l, |\text{NOM}|) = 2$. Consider $(A_2, 0), (B_2, 0) \in \mathcal{C}_{fin}$, defined as above. As above, $(A_2, 0) \sim_m^{\Diamond} (B_2, 0)$, but $A_2 \models \varphi(0)$ and $B_2 \models \neg\varphi(0)$. So claim 1 fails for these two models.

Where does the proof of the claim fail? Write $A = (A_2)^l_{<l}$ and $B = (B_2)^l_{<l}$. They work out as

$$\left.\begin{array}{rcl} A &=& (\{\hat{0}, \hat{1}, \hat{2}\}, \{(\hat{1}, \hat{2}), (\hat{2}, \hat{2})\}, \hat{V}) \\ B &=& (\{\hat{0}, \hat{1}, \hat{2}\}, \{(\hat{1}, \hat{2})\}, \hat{V}) \end{array}\right\} \quad \text{where } \hat{V}(c) = \{\hat{2}\}.$$

From this, $(A, \hat{0})$ and $(B, \hat{0})$ are indeed bisimilar, as the proof of the claim shows, but $A \models \varphi(\hat{0})$ and $B \models \neg\varphi(\hat{0})$, so they disagree on φ. Hence, (7) in the proof of the claim fails.

This does not contradict the argument that led to (7), because $(A, \hat{0})$ and $(B, \hat{0})$ are plainly not farsighted, so not in $\mathcal{C}$. So even though they are bisimilar, we cannot deduce that they agree on φ, which is known to be bisimulation invariant only over $\mathcal{C}$.

8.2.2 Basic hybrid logic

Here we extend Otto's result quoted above to basic hybrid logic. In this section, 'bisimulation' will mean '$\Diamond@$-bisimulation'.

THEOREM 8.3 *Let $\sigma = \text{PROP} \cup \text{NOM}$ be a finite hybrid signature, and $\varphi(x)$ a first-order $L(\sigma)$-formula of quantifier depth q. Let $\mathcal{C}$ be a class of pointed Kripke models for σ that is closed under bisimulation, or the class of finite models in such a class, such that φ is bisimulation invariant over $\mathcal{C}$. Then φ is equivalent over $\mathcal{C}$ to an $\mathcal{L}^{\Diamond@}(\sigma)$-formula of modal depth $\leq m = \mathsf{m}(\Diamond@, l, |\text{NOM}|)$, where m is as in definition 5.11, and $l = 2^q$.*

Proof (sketch). In the proof, all $\sim$ denote $\sim^{\Diamond@}$. Fix φ, q, l, m as above, and $(A, \mathsf{a}), (B, \mathsf{b}) \in \mathcal{C}$, and assume that $(A, \mathsf{a}) \sim_m (B, \mathsf{b})$. We will show that (A, a) and (B, b) agree on φ. The rest of the argument is then standard, as in claim 2 of the proof of theorem 6.1 and using finiteness of σ, and we will not repeat it here.

The proof outlined in figure 2 in §5 no longer works, because we cannot guarantee that the four rightmost models in it are in $\mathcal{C}$. Hence, even if they are bisimilar, they need not agree on φ.

So we take a different route, illustrated in figure 6. The items in it will be explained shortly.

$$\begin{array}{ccccccc}(A,\mathsf{a}) & \sim & (A^l,\hat{\mathsf{a}}) & \sim & (A^*,\mathsf{z}) & \sim & (\mathcal{I}(M,\boldsymbol{a}),\ell(\mathsf{z})) \\ \wr_m & & & & & & \||_{\infty q} \\ (B,\mathsf{b}) & \sim & (B^l,\hat{\mathsf{b}}) & \sim & (B^*,\mathsf{z}) & \sim & (\mathcal{I}(M,\boldsymbol{b}),r(\mathsf{z}))\end{array}$$

Figure 6: guide for bisimulation-closed classes

It will become clear that if A, B are finite then so are all models in figure 6. So if we can establish all the $\sim$ in the figure, then all models in it will be in $\mathcal{C}$, since (A,a) and (B,b) are, and $\mathcal{C}$ is bisimulation closed [in the finite]. Since φ is bisimulation invariant over $\mathcal{C}$, all the bisimilar models in the figure will then agree on φ. If we can also establish the $\equiv_{\infty q}$, then the rightmost two models will agree on φ outright, irrespective of their membership of $\mathcal{C}$. Chasing around figure 6 will then show that (A,a) and (B,b) agree on φ.

So our job is to define the models in figure 6 and verify the $\sim$ and $\equiv_{\infty q}$ in it. There are four columns of models in the figure. For the first two columns, A^l, B^l are the l-unravellings of A, B as in definition 5.1. They are finite if A, B are. We have $(A,\mathsf{a}) \sim (A^l,\hat{\mathsf{a}})$ by lemma 5.3, and similarly for B. These are absolute facts, not using the assumption that $(A,\mathsf{a}) \sim_m (B,\mathsf{b})$.

We aim to transform A^l and B^l into relational structures with a common substructure, so that corollary 4.6 applies. The third column of figure 6 is the first step towards this. From our assumption in the first column that $(A,\mathsf{a}) \sim_m (B,\mathsf{b})$, lemmas 5.10 and 5.12 give $(A^l_{<l},\hat{\mathsf{a}}) \sim (B^l_{<l},\hat{\mathsf{b}})$. *Let Z be a bisimulation witnessing this.* So

$$\mathsf{z} \stackrel{\text{def}}{=} (\hat{\mathsf{a}},\hat{\mathsf{b}}) \in Z.$$

If A, B are finite then so are $A^l_{<l}, B^l_{<l}$, and hence Z. For $c \in \textsc{Nom}$ and $(a,b),(a',b') \in Z$, define

$$c^Z = (c^{A^l_{<l}}, c^{B^l_{<l}}), \quad \text{and} \quad Z \models (a,b)R(a',b') \text{ iff } A^l_{<l} \models aRa' \text{ and } B^l_{<l} \models bRb'.$$

As Z is a $\Diamond$@-bisimulation, $c^Z \in Z$ for each c. We let $\alpha : Z \to A^l_{<l}$ and $\beta : Z \to B^l_{<l}$ be the (not necessarily surjective) projections given by $\alpha(a,b) = a$ and $\beta(a,b) = b$ for $(a,b) \in Z$. So $\alpha(\mathsf{z}) = \hat{\mathsf{a}}$ and $\beta(\mathsf{z}) = \hat{\mathsf{b}}$.

Clearly, $\operatorname{rng}\alpha \subseteq A^l_{<l} \subseteq A^l$. *Define a Kripke model A^* from A^l by replacing $\operatorname{rng}\alpha$ by Z.* So

$$\operatorname{dom}(A^*) = (\operatorname{dom}(A^l) \setminus \operatorname{rng}(\alpha)) \uplus Z.$$

If A and B are finite then so are A^l and Z, so A^* is finite in that case. Define the (surjective) projection $\alpha^* : A^* \to A^l$ as being α on Z and the identity elsewhere. For each $t, u \in A^*$, define:

- $A^*, t \models p$ iff $A^l, \alpha^*(t) \models p$, for each $p \in \sigma$,
- $A^* \models tRu$ iff $\{t,u\} \subseteq Z$ and $Z \models tRu$, or $\{t,u\} \not\subseteq Z$ and $A^l \models \alpha^*(t)R\alpha^*(u)$.

Then $c^{A^*} = c^Z$ for each nominal c, and it can be checked that α^* is a bisimulation between A^* and A^l witnessing $(A^l,\hat{\mathsf{a}}) \sim (A^*,\mathsf{z})$, as stated in figure 6.

Define B^* similarly, by replacing $\operatorname{rng}\beta$ in B^l by Z. Again, if A, B are finite then so is B^*, and $(B^l, \hat{\mathsf{b}}) \sim (B^*, \mathsf{z})$ as stated in figure 6.

For the final column of figure 6, let

$$Z_0 = \{\mathsf{z}\} \cup \{c^Z : c \in \text{NOM}\} \subseteq Z, \quad \text{and} \quad \boldsymbol{z} = (z : z \in Z_0).$$

Let K be the relational signature obtained from $L(\sigma)$ by deleting all constants (nominals) and adding new unary relation symbols P_z $(z \in Z_0)$. Let A^{*K} be the K-structure obtained from A^* by:

- keeping the same domain and the same interpretations of unary relation symbols in $L(\sigma)$ (ie. the P for $p \in \text{PROP}$),
- deleting all R-arrows into Z_0: so $A^{*K} \models tRu$ iff $A^* \models tRu$ and $u \notin Z_0$,
- interpreting P_z as $\{t \in A^{*K} : A^* \models tRz\}$, for each $z \in Z_0$.

Define B^{*K} from B^* in the same way. Then A^{*K} and B^{*K} have a common K-substructure Z^K with domain Z, and it can be checked that

$$\mathcal{N}_l^{A^{*K}}(\boldsymbol{z}) \subseteq Z^K \quad \text{and} \quad \mathcal{N}_l^{B^{*K}}(\boldsymbol{z}) \subseteq Z^K. \tag{9}$$

Let $M = A^{*K} + B^{*K}$ (disjoint union), and let $\ell : A^{*K} \hookrightarrow M$ and $r : B^{*K} \hookrightarrow M$ be the usual K-embeddings. If A, B are finite then so are A^{*K} and B^{*K}, and hence M. Let

$$\boldsymbol{a} = \ell(\boldsymbol{z}) \quad \text{and} \quad \boldsymbol{b} = r(\boldsymbol{z}).$$

By (9) and corollary 4.6, we obtain

$$(M, \boldsymbol{a}) \equiv_{\infty q} (M, \boldsymbol{b}). \tag{10}$$

We now get $L(\sigma)$ back, by an interpretation $\mathcal{I}$ as in definition 5.6, but with parameters Z_0 this time, with $\mathcal{I}(xRy)$ defined as $xRy \vee \bigvee_{z \in Z_0}(P_z(x) \wedge y = v_z)$, and $\mathcal{I}(\alpha(x_1, \ldots, x_n, c_1, \ldots, c_{n'}))$ defined by substituting $v_{c_i^Z}$ for y_i in $\mathcal{I}(\alpha(x_1, \ldots, x_n, y_1, \ldots, y_{n'}))$ for $i = 1, \ldots, n'$.

The $L(\sigma)$-structures $\mathcal{I}(A^{*K}, \boldsymbol{z})$ and $\mathcal{I}(B^{*K}, \boldsymbol{z})$ exist and are A^* and B^*, respectively. The $L(\sigma)$-structures $\mathcal{I}(M, \boldsymbol{a})$ and $\mathcal{I}(M, \boldsymbol{b})$ also exist, and are finite if A, B are, since they have the same domain as M. By lemma 4.7(2) as in lemma 5.7, the maps

$$\ell : A^* \hookrightarrow \mathcal{I}(M, \boldsymbol{a}), \qquad r : B^* \hookrightarrow \mathcal{I}(M, \boldsymbol{b})$$

are $L(\sigma)$-embeddings and in fact bisimulations. This justifies the rightmost two $\sim$ in figure 6.

Lastly, by (10) and lemma 4.7(3), $(\mathcal{I}(M, \boldsymbol{a}), \ell(\mathsf{z})) \equiv_{\infty q} (\mathcal{I}(M, \boldsymbol{b}), r(\mathsf{z}))$, as stated on the far right of figure 6. This completes our justification of figure 6. □

We make some remarks on this theorem. Example 7.1 showed that its modal depth bound is optimal in general, though it might be bettered for particular $\mathcal{C}$ and/or φ. Example 8.4 below shows that the assumption that NOM is finite is necessary. However, the theorem holds for arbitrary PROP and finite NOM: we leave this as an exercise. It would be sufficient to prove the theorem for the class of all [finite] pointed Kripke models for σ that agree on φ with every [finite] bisimilar model, because it meets the criteria of the theorem and contains each $\mathcal{C}$ in the theorem.

Theorem 8.3 implies theorem 6.1 in the case $\star = \Diamond@$ and for finite σ. But it cannot replace theorem 6.1. As example 8.4 below shows, it fails for infinite NOM. More importantly, as example 8.2

showed, it fails for $\star = \Diamond$. The reason is that in that case, not all nominals need be instantiated in Z, and their instantiations in A, B, and hence in A^*, B^*, may have different properties. For example, a nominal might name a reflexive point in A but not in B, in which case, even if we manage to define the models in figure 6, the $\equiv_{\infty q}$ will fail. This is the basis of example 8.2.

EXAMPLE 8.4 We exhibit a hybrid signature σ, a class $\mathcal{C}$ of pointed Kripke models for σ that is closed under bisimulation, and a first-order $L(\emptyset)$-formula $\varphi(x)$ that is bisimulation invariant over $\mathcal{C}$, but not equivalent even over its subclass $\mathcal{C}_{\mathit{fin}}$ of finite models to any $\mathcal{L}^{\Diamond @}(\tau)$-formula for any $\tau \subsetneq \sigma$.

Let σ be a nonempty hybrid signature with no propositional atoms: so $\sigma = \textsc{Nom}$. As in example 8.2, we allow arbitrary disjunctions of hybrid formulas. A (pointed Kripke) model (A, a) for σ is said to be *named* if $A, a \models \bigvee \sigma$ — some nominal names the point a of the model. Let $\mathcal{C}$ be the class of all named models. Since $\bigvee \sigma$ is bisimulation invariant, $\mathcal{C}$ is closed under bisimulation.

Let $\varphi(x)$ be the $L(\emptyset)$-formula xRx. Over named models, φ is equivalent to $\bigvee_{c \in \sigma}(c \wedge \Diamond c)$, so φ is bisimulation invariant over $\mathcal{C}$.

Now let $\tau \subsetneq \sigma$. We define two Kripke models for σ. They are $A = (\{0,1\}, \{(1,1)\}, V_\tau)$ and $B = (\{0,1,2\}, \{(1,2),(2,2)\}, V_\tau)$, where for each $c \in \sigma$,

$$V_\tau(c) = \begin{cases} \{0\}, & \text{if } c \in \tau, \\ \{1\}, & \text{otherwise.} \end{cases}$$

Then $(A,1), (B,1) \in \mathcal{C}_{\mathit{fin}}$, because they are finite and any $c \in \sigma \setminus \tau$ names 1 in both A, B. Evidently, $\{(0,0),(1,1),(1,2)\}$ is a bisimulation between $A \restriction \tau$ and $B \restriction \tau$. So $(A \restriction \tau, 1) \sim^{\Diamond @} (B \restriction \tau, 1)$, and by fact 3.3, these models, and hence also $(A,1)$ and $(B,1)$, agree on $\mathcal{L}^{\Diamond @}(\tau)$-formulas. But $A \models \varphi(1)$ and $B \models \neg\varphi(1)$, so they do not agree on φ. Hence, φ, an $L(\tau)$-formula, is not equivalent over $\mathcal{C}_{\mathit{fin}}$ to any $\mathcal{L}^{\Diamond @}(\tau)$-formula.

We make some remarks on this example. For $\tau \subsetneq \sigma$, the class $\{(A \restriction \tau, a) : (A, a) \in \mathcal{C}\}$ is the class of *all* pointed Kripke models for τ, so is bisimulation closed. The example showed that φ is not bisimulation invariant over it (or even over the class of finite models in it). Contrast this with lemma 4.1. If σ is finite then $\mathcal{C}$ is elementary. And if σ is infinite, then φ is not equivalent over $\mathcal{C}_{\mathit{fin}}$ to any $\mathcal{L}^{\Diamond @}(\sigma)$-formula, because any such formula would be an $\mathcal{L}^{\Diamond @}(\tau)$-formula for some finite $\tau \subsetneq \sigma$, contradicting the example. So the assumption in theorem 8.3 that $\textsc{Nom}$ is finite is necessary. The theorem is not at all as robust as theorem 6.1.

8.3 Temporal hybrid logic and/or with universal modality

So-called 'temporal' logic considers the converse diamond $\Diamond^{-1}$, with semantics $M, w \models \Diamond^{-1}\psi$ iff $M, u \models \psi$ for some $u \in M$ with $M \models uRw$. The universal modality $\forall$ has semantics $M, w \models \forall\psi$ iff $M, u \models \psi$ for every $u \in M$. One can extend proto- and basic hybrid logic by either or both of these modalities. In each case, the notion of bisimulation is amended in the expected way, giving *two-way* bisimulations, *global* bisimulations, or both [17, §4.2], [18, §4], [13, §5.1].

For modal logic, Otto [17, 18] gave a bisimulation characterisation theorem both classically and in the finite for the combined temporal-$\forall$ extension. Whether this can be done in the finite for proto- or basic hybrid logic extended by $\Diamond^{-1}$ and/or $\forall$ is an open question, though the classical case holds by van Benthem's original argument.

Our proof of theorem 6.1 does not work in the temporal case. One can upgrade the unravelling A^l to allow for $\Diamond^{-1}$, but R- and R^{-1}-arrows from long paths to paths named by nominals can now be followed backwards by bisimulations. Lemma 5.7 may fail, and in fact, locality itself may fail: $(A^l, \hat{a})$ and $(A^l_{<l}, \hat{a})$ may disagree even on simple hybrid formulas such as $\Box\Diamond\top$. It seems that the 'neighbourhood' $A^l_{<l}$ is no longer appropriate, but finding an alternative that works would probably require methods beyond this paper. Ditto for $\forall$.

9 Conclusion

We have proved a characterisation theorem for proto-hybrid logic (modal logic with nominals) and for basic hybrid logic (modal logic with nominals and @), uniformly over arbitrary and finite Kripke models, and with optimal modal depth bounds. We also showed that for basic hybrid logic, but not for proto-hybrid logic, the theorem extends to arbitrary bisimulation-closed classes and to the finite members of such classes.

These results are profoundly incremental, they fly in the face of the modern trend to generality in the subject (e.g., [3, 21, 5]), and the classical cases (excepting perhaps bisimulation-closed classes and the modal depth bounds) are already known. Still, the finite-models cases may fill a narrow but striking gap in knowledge. In particular, proto-hybrid logic is a rather small and obvious extension of modal logic — just add nominals. So whether it is, like modal logic, characterised by invariance under modal bisimulations in the finite seems a basic question to which modal and hybrid logicians ought to have an answer. Now they do.

Bisimulation characterisation theorems over finite models are still relatively rare. Whilst they do sometimes follow from very general results, such as [21], at other times they can be challenging or even impossible to achieve [11]. So it may be of some value to add a few more for hybrid logics, as we have done here. And the methods we have used, in particular unravelling compatibly with nominals, the use of interpretations, and proposition 4.5, may be helpful elsewhere. However, readers should feel free to alter these methods — most things in this paper can be done differently.

One might ask about characterisation theorems in the finite for more powerful hybrid logics. We mentioned the cases of temporal operators and the universal modality as open problems, but there are many more, and the picture there may not be so rosy. Another problem is to prove characterisation theorems over particular classes of finite models. For example, results for modal logic over finite transitive models are known [11], and it may be interesting to extend them to hybrid logics. Finally, we have mentioned characterisation theorems that hold classically but not in the finite. It might be interesting to find characterisation theorems 'in the wild' that hold in the finite but not classically. Certainly there are trivial illustrative examples (not using bisimulations). Every first-order formula $\varphi(x)$ is equivalent to a propositional boolean formula over the class of pointed Kripke models based on finite irreflexive dense linear orders, because each such order is just a solitary irreflexive point. This fails for the class of all irreflexive dense linear orders.

One final remark: although the result proved in [21] is extremely general, apparently our results here cannot be derived from it, because as it stands, nominals are not admitted. However, coalgebraic semantics has been developed for hybrid logic [16, 20], and [21] may well be extended to hybrid logic in due course.

References

[1] S. Abramsky and D. Marsden, *Comonadic semantics for hybrid logic*, 47th International Symposium on Mathematical Foundations of Computer Science (MFCS 2022) (Dagstuhl, Germany) (S. Szeider, R. Ga-

nian, and A. Silva, eds.), Leibniz International Proceedings in Informatics (LIPIcs), vol. 241, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022, pp. 7:1–7:14.

[2] C. Areces, P. Blackburn, and M. Marx, *Hybrid logics: Characterization, interpolation and complexity*, J. Symbolic Logic **66** (2001), 977–1010.

[3] C. Areces, F. Carreiro, and S. Figueira, *Characterization, definability and separation via saturated models*, Theoretical Comp. Sci. **537** (2014), 72–86.

[4] C. Areces and B. ten Cate, *Hybrid logics*, in van Benthem et al. [8], pp. 821–868.

[5] G. Badia, D. Gaina, A. Knapp, T. Kowalski, and M. Wirsing, *A modular bisimulation characterisation for fragments of hybrid logic*, Bull. Symbolic Logic (2025), 1–24, https://doi.org/10.1017/bsl.2025.9.

[6] J. van Benthem, *Modal correspondence theory*, Ph.D. thesis, Mathematical Institute, University of Amsterdam, 1976.

[7] ———, *Modal logic and classical logic*, Bibliopolis, Naples, 1985.

[8] J. van Benthem, P. Blackburn, and F. Wolter (eds.), *Handbook of modal logic*, Studies in Logic and Practical Reasoning, vol. 3, Elsevier, Amsterdam, 2007.

[9] P. Blackburn and J. van Benthem, *Modal logic: a semantic perspective*, in van Benthem et al. [8], pp. 1–84.

[10] B. ten Cate, *Model theory for extended modal languages*, Ph.D. thesis, University of Amsterdam, 2005, ILLC Dissertation Series DS-2005-01.

[11] A. Dawar and M. Otto, *Modal characterisation theorems over special classes of frames*, Ann. Pure. Appl. Logic **161** (2009), 1–42.

[12] H. Gaifman, *On local and non-local properties*, Proc. Herbrand Symposium (J. Stern, ed.), Studies in Logic and the Foundations of Mathematics, vol. 107, North-Holland, 1982, pp. 105–135.

[13] V. Goranko and M. Otto, *Model theory of modal logic*, in van Benthem et al. [8], pp. 249–329.

[14] W. Hodges, *Model theory*, Encyclopedia of mathematics and its applications, vol. 42, Cambridge University Press, 1993.

[15] I. Hodkinson and H. Tahiri, *A bisimulation characterization theorem for hybrid logic with the current-state binder*, Rev. Symbolic Logic **3** (2010), 247–261.

[16] R. Myers, D. Pattinson, and L. Schröder, *Coalgebraic hybrid logic*, Proc. FOSSACS 2009 (L. de Alfaro, ed.), Lecture Notes in Computer Science, vol. 5504, Springer, 2009, pp. 137–151.

[17] M. Otto, *Elementary proof of the van Benthem-Rosen characterisation theorem*, Tech. Report 2342, Department of Mathematics, Darmstadt University of Technology, 2004.

[18] ———, *Bisimulation invariance and finite models*, Logic Colloquium '02 (Z. Chatzidakis, P. Koepke, and W. Pohlers, eds.), Lecture Notes in Logic, Cambridge University Press, 2006, p. 276–298.

[19] E. Rosen, *Modal logic over finite structures*, J. Logic, Language and Information **6** (1997), 427–439.

[20] L. Schröder and D. Pattinson, *Named models in coalgebraic hybrid logic*, 27th International Symposium on Theoretical Aspects of Computer Science - STACS 2010 (J.-Y. Marion and T. Schwentick, eds.), Leibniz International Proceedings in Informatics, 2010, pp. 645–656.

[21] L. Schröder, D. Pattinson, and T. Litak, *A van Benthem/Rosen theorem for coalgebraic predicate logic*, J. Logic Computat. **27** (2017), 749–773.