On the expressive power of inquisitive epistemic logic

Ivano Ciardelli and Martin Otto

November 2024

Abstract

Inquisitive modal logic, INQML, in its epistemic incarnation, extends standard epistemic logic to capture not just the information that agents have, but also the questions that they are interested in. We use the natural notion of bisimulation equivalence in the setting of INQML, as introduced in [10], to characterise the expressiveness of INQML as the bisimulation invariant fragment of first-order logic over natural classes of two-sorted first-order structures that arise as relational encodings of inquisitive epistemic (S5-like) models. The non-elementary nature of these classes crucially requires non-classical model-theoretic methods for the analysis of first-order expressiveness, irrespective of whether we aim for characterisations in the sense of classical or of finite model theory.

Contents

1 Introduction

Inquisitive logic [11, 13] is a framework that allows us to extend various systems of classical logic with formulae that represent questions. Thus, in the propositional setting, in addition to the formula p representing the statement that p, we also have, e.g., a formula $?p$ representing the question *whether p*. In the first-order setting, in addition to the formula $\forall x Px$ representing the statement that all objects are P, we also have, e.g., a formula $\forall x ?Px$, representing the question *which objects are P*. Technically, this extension is achieved by using standard models for intensional logic, but interpreting formulae not relative to single possible worlds, but relative to sets of worlds, called *information states*. As such, inquisitive logics fit within the family of logics based on versions of *team semantics*, where formulae are evaluated relative to sets of (standard) semantic parameters.[1]

Inquisitive epistemic logic [12, 2, 4, 18, 32] is an extension of standard epistemic logic which, building on inquisitive logic, allows us to reason not only about the knowledge that agents have, but also about the issues they are interested in. In this logic we can, for instance, embed the question $?p$ under suitable inquisitive modalities, resulting in modal statements which express that a certain agent *knows whether p*, or that a certain agent *wonders whether p*. Furthermore, in a dynamic extension of the logic, we can model not only the communicative effect of publicly making statements (as in public announcement logic [31]), but also the effect of publicly asking questions, describing in particular how both kinds of moves affect the issues agents are interested in.

From a technical point of view, inquisitive epistemic logic is a special case of inquisitive (multi-)modal logic. As in Kripke semantics, models for this logic assign to each world w and agent a a set $\Sigma_a(w)$ of successors, which inquisitive modalities then quantify over; however, differently from Kripke semantics, these successors are not worlds, but rather information states, i.e., sets of worlds. This means that, in the inquisitive setting, modalities involve a restricted kind of second-order quantification. Within this general modal setup, inquisitive epistemic logic arises as a special case when we impose $S5$-like constraints on the modal frame, motivated as usual by the assumptions of factivity of knowledge and full introspection of agents.

Our topic in this paper is the expressive power of inquisitive epistemic logic, viewed as a language for describing properties of worlds as well as properties of information states in the setting of inquisitive epistemic models. In particular, we are interested in comparing the expressive power of inquisitive epistemic logic with that of classical first-order predicate logic FO interpreted over (suitable relational encodings of) inquisitive epistemic models.

A crucial role in this study is played by a notion of *bisimilarity* for inquisitive modal models, identified in our previous paper [10]. For inquisitive modal logic interpreted over arbitrary (or arbitrary finite) models, a van Benthem/Rosen characterization result was proved in [10].

---

[1]Team semantics have been used to interpret independence-friendly logic [23, 24] and many variants of dependence and independence logic (see, among others, [34, 35, 1, 17, 37, 38]).

Theorem 1.1 ([10])**.** *Inquisitive modal logic can be characterised as the bisimulation-invariant fragment of first-order logic* FO *over some natural classes of (finite or arbitrary) relational inquisitive models.*

In this paper, we will prove an analogous result for the specific setting of inquisitive epistemic logic, where models are required to obey $S5$-like constrains.

Theorem 1.2 (main theorem)**.** *Inquisitive epistemic logic (in a multi-agent setting) can be characterised as the bisimulation-invariant fragment of* FO *over some natural classes of (finite or arbitrary) relational inquisitive epistemic models.*

Extending Theorem 1.1 to the inquisitive epistemic case is a challenging enterprise: the reason is that the proof in [10] crucially relies on a step in which we unfold a model into a bisimilar locally tree-like model. That step, however, is incompatible with the $S5$-like requirements on inquisitive epistemic models, which rule out simple tree-like structures. We here show that locally tree-like behaviour can be achieved for relational structures in which suitable bisimilar companions are first-order interpretable. Our analysis of the bisimulation types of epistemic inquisitive models and their relational encodings is then based on an analysis of these representations that supports the central upgrading argument linking finite levels of bisimulation equivalence to levels of elementary equivalence between the interpreted models. Equivalences between the locally tree-like representations are assessed through a combination of classical back-and-forth techniques that capture the expressive power of first-order as well as of restricted forms of monadic second-order logic. In these aspects the current paper technically builds on but also considerably exceeds the treatment in [10].

Besides its technical interest, Theorem 1.2 is a significant result from a conceptual point of view. It entails that inquisitive epistemic logic is a natural choice for a language designed to express bisimulation-invariant properties of pointed inquisitive epistemic models. On the one hand, this language allows us to express *only* properties that are bisimulation invariant; this is desirable if one views bisimilar points as distinct representations of what is essentially the same state of affairs, so that aspects of a model which are not bisimulation invariant are spurious artefacts of the representation that our logical language should be insensitive to. On the other hand, as the theorem states, this language allows us to express *all* the bisimulation-invariant properties that are expressible in first-order predicate logic, thus enjoying a form of expressive completeness.

The paper is structured as follows: Section 2 introduces all the relevant background for our contribution; Section 3 contains the steps leading up to the proof of Theorem 1.2; Section 4 summarises our results and outlines some directions for further work.

4

2 Background

In this section we review some background for our contribution. First, we briefly introduce inquisitive epistemic logic [12, 4], as well as the associated notion of bisimulation [10]. We also show how to encode models of inquisitive epistemic models as relational structures, following [10], and recall how the expressive power of first-order logic and monadic second-order logic can be characterised by means of Ehrenfeucht–Fraïssé games.

2.1 Inquisitive epistemic models

Let us fix a *signature* for our logic, consisting of a set $\mathcal{A}$ of agents and a set $\mathcal{P}$ of atomic propositions. Relative to such a signature, the notions of inquisitive epistemic frames and models are defined as follows.

Definition 2.1. [inquisitive epistemic models]
An *inquisitive epistemic frame* is a pair $\mathbb{F} = (W, (\Sigma_a)_{a \in \mathcal{A}})$, where W is a non-empty set whose elements are called *worlds*, and for each $a \in \mathcal{A}$, the map $\Sigma_a : W \to \wp\wp(W)$ assigns to each world w a set of sets of worlds $\Sigma_a(w)$ in accordance with the following constraints, where $\sigma_a(w) := \bigcup \Sigma_a(w)$. For all $w, v \in W$, $s, t \subseteq W$, and $a \in \mathcal{A}$ we must have that:

- $\Sigma_a(w)$ is non-empty;

- if $s \in \Sigma_a(w)$ and $t \subseteq s$ then $t \in \Sigma_a(w)$ (downward closure);

- $w \in \sigma_a(w)$ (factivity);

- if $v \in \sigma_a(w)$ then $\Sigma_a(v) = \Sigma_a(w)$ (full introspection).

An inquisitive epistemic model consists of an inquisitive epistemic frame together with a propositional assignment $V : \mathcal{P} \to \wp(W)$.

A set of worlds $s \subseteq W$ is referred to as an *information state*; a non-empty set of information states which is downward closed (i.e., such that whenever it contains an information state s, it contains every subset of s) is called an *inquisitive state*. Thus, the above definition says that each map Σ_a assigns to a world w an inquisitive state $\Sigma_a(w)$, in accordance with factivity and introspection.

In the intended interpretation for inquisitive epistemic models [12, 2], the map Σ_a is taken to describe both the *knowledge* and the *issues* of agent a. The agent's knowledge state at w, $\sigma_a(w) = \bigcup \Sigma_a(w)$, consists of all the worlds that are compatible with what the agent knows. The agent's inquisitive state at w, $\Sigma_a(w)$, consists of all those information states where the agent's issues are settled. This interpretation motivates the first two constraints on $\Sigma_a(w)$: $\Sigma_a(w)$ should be downward closed, since if the agent's issues are settled by an information state s, they will also be settled by any stronger information state $t \subseteq s$; furthermore, the agent's issues will always be trivially settled by the empty information state, $\emptyset$, which represents the state of inconsistent information. The factivity constraint amounts, as usual, to the requirement that the agent's

5

knowledge at a world be truthful. The introspection constraint requires agents to have full knowledge of their own inquisitive state: at each world w, only worlds v where the agent's state is the same as in w will be compatible with the agent's knowledge.

Note that with an inquisitive epistemic frame $\mathbb{F}$ we can always associate a standard Kripke frame $\mathfrak{K}(\mathbb{F})$ having the same set of worlds and accessibility relations defined by $R_a := \{(v, w) : v \in \sigma_a(w)\}$. Similarly, with an inquisitive epistemic model $\mathbb{M}$ we can associate a standard Kripke model $\mathfrak{K}(\mathbb{M})$. It is easy to verify that the Kripke frame associated with an inquisitive epistemic frame is always an $S5$ frame, i.e., all the accessibility relations R_a are equivalence relations on W. We denote as $[w]_a$ the equivalence class of world w w.r.t. R_a, and by $W/R_a = \{[w]_a : w \in W\}$ the quotient consisting of these equivalence classes, which form a partition of the set W of possible worlds, reflecting the possible knowledge states of agent a. The introspection condition ensures that the map Σ_a takes a constant value on each of these equivalence classes.

2.2 Inquisitive modal logic

The syntax of inquisitive modal logic INQML is given by:

$$\varphi \ ::= \ p \,|\, \bot \,|\, (\varphi \wedge \varphi) \,|\, (\varphi \rightarrow \varphi) \,|\, (\varphi \mathbin{\vee\mkern-11mu\vee} \varphi) \,|\, \Box_a \varphi \,|\, \boxplus_a \varphi,$$

where $p \in \mathcal{P}$ stands for a basic proposition, and $a \in \mathcal{A}$ is the labelling of inquisitive modalities by agents. Nesting depth w.r.t. the modal operators $\Box_a$ and $\boxplus_a$ (combined and across all $a \in \mathcal{A}$) is syntactically defined as usual, and for $n \in \mathbb{N}$ we let $\textsc{InqML}_n \subseteq \textsc{InqML}$ stand for the restriction to nesting depth up to n. The mono-modal case, which is the focus of [10], may be regarded as a special case, where the label for a single agent is omitted as irrelevant.

We treat negation and disjunction as defined connectives (syntactic shorthands) according to:

$$\neg \varphi := \varphi \rightarrow \bot \qquad\qquad \varphi \vee \psi := \neg(\neg \varphi \wedge \neg \psi)$$

With these abbreviations, the above syntax can be seen as including the syntax of standard modal logic, based on Boolean connectives and the modalities $\Box_a$. In particular, $\textsc{InqML}_0$ includes all formulae standard propositional logic.

In addition to standard connectives, the language contains *inquisitive disjunction* $\mathbin{\vee\mkern-11mu\vee}$, which is interpreted as a question-forming (inclusive) disjunction: for instance, whereas the formula $p \vee \neg p$ stands for the (tautological) statement that p is either true or false, the formula $p \mathbin{\vee\mkern-11mu\vee} \neg p$ (abbreviated as $?p$) stands for the question *whether* p is true or false. Equally importantly, the language contains two modalities for each agent, which allow us to embed both statements and questions. Both kinds, $\Box_a$ and $\boxplus_a$ coincide with a standard Kripke box modality $\Box_a$ when applied to statements, but differ in application to questions.[2]

---

[2]Modalities $\Box_a$ and $\boxplus_a$ are usually read as "box" and "window". In the epistemic setting (e.g., in [12]), they are denoted K_a and E_a, and read as "know" and "entertain" respectively.

In line with the inquisitive approach, the semantics for INQML is not given by a recursive definition of *truth* relative to a possible world, but rather by a recursive definition of a notion of *support* relative to an information state, i.e., a set of worlds.

Definition 2.2. [semantics of INQML]
Let $\mathbb{M} = (W, (\Sigma_a), V)$ be an inquisitive epistemic model, $s \subseteq W$:

- $\mathbb{M}, s \models p \iff s \subseteq V(p)$

- $\mathbb{M}, s \models \bot \iff s = \emptyset$

- $\mathbb{M}, s \models \varphi \wedge \psi \iff \mathbb{M}, s \models \varphi$ and $\mathbb{M}, s \models \psi$

- $\mathbb{M}, s \models \varphi \rightarrow \psi \iff$ for all $t \subseteq s : \mathbb{M}, t \models \varphi \Rightarrow \mathbb{M}, t \models \psi$

- $\mathbb{M}, s \models \varphi \vee\!\!\!\vee \psi \iff \mathbb{M}, s \models \varphi$ or $\mathbb{M}, s \models \psi$

- $\mathbb{M}, s \models \Box_a \varphi \iff$ for all $w \in s : \mathbb{M}, \sigma_a(w) \models \varphi$

- $\mathbb{M}, s \models \boxplus_a \varphi \iff$ for all $w \in s$, for all $t \in \Sigma_a(w) : \mathbb{M}, t \models \varphi$

The following two properties hold generally in INQML.

Proposition 2.3.

- *persistency: if $\mathbb{M}, s \models \varphi$ and $t \subseteq s$, then $\mathbb{M}, t \models \varphi$;*

- *empty state property (semantic ex-falso): $\mathbb{M}, \emptyset \models \varphi$ for all $\varphi \in$ INQML.*

The first item says that support is preserved as information increases, i.e., as we move from a state to an extension of it. The second item says that the inconsistent information state vacuously supports every formula; this can be seen as a semantic counterpart of the *ex falso quodlibet* principle. These principles imply that the support set $[\varphi]_\mathbb{M} := \{s \subseteq W : \mathbb{M}, s \models \varphi\}$ of a formula is downward closed and non-empty, i.e., it is an inquisitive state.

Although the primary notion in INQML is support at an information state, truth at a world is retrieved as a derivative notion.

Definition 2.4. [truth]
φ is *true* at a world w in a model $\mathbb{M}$, denoted $\mathbb{M}, w \models \varphi$, in case $\mathbb{M}, \{w\} \models \varphi$.

Spelling out Definition 2.2 in the special case of singleton states, we see that standard connectives have the usual truth-conditional behaviour. For modal formulae, we find the following truth-conditions:

- $\mathbb{M}, w \models \Box_a \varphi \iff \mathbb{M}, \sigma_a(w) \models \varphi$

- $\mathbb{M}, w \models \boxplus_a \varphi \iff \forall t \in \Sigma_a(w) : \mathbb{M}, t \models \varphi$

Notice that truth in InqML cannot be given a direct recursive definition, as the truth conditions for modal formulae $\Box_a\varphi$ and $\boxplus_a\varphi$ depend on the support conditions for φ — not just on its truth conditions.

For many formulae, support at a state just boils down to truth at each world in the state. We refer to these formulae as *truth-conditional*.[3]

Definition 2.5. [truth-conditional formulae]
We say that a formula φ is *truth-conditional* if for all models $\mathbb{M}$ and information states s: $\mathbb{M}, s \models \varphi \iff \mathbb{M}, w \models \varphi$ for all $w \in s$.

Following [4], we view truth-conditional formulae as statements, and non-truth-conditional formulae as questions. It is immediate to verify that atomic formulae, $\bot$, and all formulae of the form $\Box_a\varphi$ and $\boxplus_a\varphi$ are truth-conditional. Furthermore, the class of truth-conditional formulae is closed under all connectives except for $\lor\!\!\!\lor$. This means that if all occurrences of $\lor\!\!\!\lor$ in φ are within the scope of a modality, then φ is truth-conditional. Using this fact, it is easy to see that all formulae of standard modal logic, i.e., formulae which do not contain occurrences of $\lor\!\!\!\lor$ or $\boxplus_a$, receive exactly the same truth-conditions as in standard modal logic.

Proposition 2.6. *If φ is a formula not containing $\lor\!\!\!\lor$ or $\boxplus_a$, then $\mathbb{M}, w \models \varphi \iff \mathfrak{K}(\mathbb{M}), w \models \varphi$ in standard Kripke semantics.*

As long as questions are not around, the modality $\boxplus_a$ also coincides with $\Box_a$, and with the standard box modality. That is, if φ is truth-conditional, then

$$M, w \models \Box_a\varphi \iff M, w \models \boxplus_a\varphi \iff M, v \models \varphi \text{ for all } v \in \sigma_a(w).$$

By contrast, the two modalities come apart when they are applied to questions. Consider, for instance, the formulae $\Box_a?p$ and $\boxplus_a?p$. According to the above support definition, the formula $?p$ (i.e., $p \lor\!\!\!\lor \neg p$) is supported by an information state s just in case p has the same truth value at all worlds in s—in other words, just in case the information available in s settles the question whether p. Thus, according to the truth conditions for modal formulae, $\Box_a?p$ is true at a world w iff the knowledge state $\sigma_a(w) = [w]_a$ of agent a, settles the question $?p$. Thus, $\Box_a?p$ expresses the fact that the agent *knows whether p*. By contrast, $\boxplus_a?p$ is true iff any information state $t \in \Sigma_a(w)$ (i.e., any state that settles the agent's issues) also settles $?p$; thus, roughly speaking, $\boxplus_a?p$ expresses that settling the question whether p is part of the agent's goals, i.e., the agent *is interested in whether p* (see [12, 4] for more detailed discussion).

Inquisitive modal formulae can be interpreted both relative to information states and (derivatively) relative to worlds. They can thus be seen both as expressing properties of state-pointed models, and as expressing properties of world-pointed models. In general, by a property of world-pointed or state-pointed models we mean a class of such pointed models. A property is *definable* in InqML if it is the class of models where some formula $\varphi \in$ InqML is true (in

---

[3]In team semantic terminology (e.g., [34, 38]), truth-conditional formulae are called *flat*.

the world-pointed case) or supported (in the state-pointed case).[4] For the case of state-pointed models, Proposition 2.3 implies that any definable property will be an *inquisitive property*, in the sense of the following definition.

Definition 2.7. [inquisitive properties] A property $\mathcal{K}$ of state-pointed models is said to be *inquisitive* if it is downward closed with respect to the information state (i.e., $(\mathbb{M}, s) \in \mathcal{K}$ implies $(\mathbb{M}, t) \in \mathcal{K}$ for all $t \subseteq s$) and satisfies the empty state property (i.e., $(\mathbb{M}, \emptyset) \in \mathcal{K}$ for all $\mathbb{M}$).

In the rest of the paper, when dealing with properties of state-pointed models, we can thus restrict our attention to inquisitive properties. We now turn to the task of giving a precise characterization of which world-properties and inquisitive state-properties are definable in INQML.

2.3 Inquisitive bisimulation

An inquisitive modal model can be seen as a structure with two closely linked sorts of entities: worlds and information states. Information states s are sets of worlds, determined by their elements; worlds w carry, besides the propositional assignment, inquisitive assignments $\Sigma_a(w)$, which are sets of of information states. The natural notion of bisimilarity, and its finite approximations of n-bisimilarity for $n \in \mathbb{N}$, capture notions of equivalence based on a step-wise probing of these links. We here focus on a corresponding game-based formulation for levels of bisimulation equivalence in terms of winning strategies in two-player bisimulation games, which highlight the dynamic nature of the "probing" of behavioural equivalence. (The alternative formalisation in terms of back-and-forth systems is easily seen to be equivalent as usual, cf. [10] for more detail.)

The game is played by two players, **I** and **II**, who act as challenger and defender of a similarity claim involving a pair of worlds w and w' or information states s and s' over two models $\mathbb{M} = (W, (\Sigma_a), V)$ and $\mathbb{M}' = (W', (\Sigma'_a), V')$. We distinguish world-positions $(w, w') \in W \times W'$ and state-positions as $(s, s') \in \wp(W) \times \wp(W')$, which match two worlds or two information states, respectively. The game alternates between world-positions and state-positions, reflecting the links between information states and worlds in $\mathbb{M}$ and $\mathbb{M}'$, with a view to testing or probing similarities or distinguishing features for the currently matched pairs, as follows. In a world-position (w, w'), **I** chooses an agent $a \in \mathcal{A}$ and an information state in the inquisitive state associated to one of these worlds, i.e. either some $s \in \Sigma_a(w)$ or some $s' \in \Sigma'_a(w')$, and **II** must respond by choosing an information state for the same agent a on the opposite side; this exchange results in a state-position (s, s'). In a state-position (s, s'), **I** chooses a world in either state, i.e. either some $w \in s$ or some $w' \in s'$, and **II** must respond likewise on the opposite side; this results in a world-position (w, w').

A round of the game consists of four moves leading from a world-position to another. In the bounded version of the game, the number of rounds is fixed

---

[4]In connection with INQML-definability by single formulae we implicitly restrict attention to finite modal signatures, i.e. $\mathcal{C}$ here is, w.l.o.g. specified in a signature with finite $\mathcal{A}$ and $\mathcal{P}$.

in advance, while the game is allowed to go on indefinitely in the unbounded version. Either player loses when stuck for a move, which in epistemic models can only happen in state positions involving the empty information state. The game also ends, with a loss for **II**, in any world-position (w, w') in which w and w' disagree on the truth of some $p \in \mathcal{P}$. All other plays, and in particular infinite runs of the unbounded game, are won by **II**. The following definition applies to arbitrary inquisitive modal models (which do not necessarily satisfy factivity and introspection), but here we are specifically interested in its application to the case of inquisitive epistemic models.

Definition 2.8. [bisimulation equivalence]
Two world-pointed models $\mathbb{M}, w$ and $\mathbb{M}', w'$ are *n-bisimilar*, $\mathbb{M}, w \sim^n \mathbb{M}', w'$, if **II** has a winning strategy in the n-round game starting from (w, w'). $\mathbb{M}, w$ and $\mathbb{M}', w'$ are *bisimilar*, denoted $\mathbb{M}, w \sim \mathbb{M}', w'$, if **II** has a winning strategy in the unbounded game starting from (w, w'). Two state-pointed models $\mathbb{M}, s$ and $\mathbb{M}', s'$ are $(n\text{-})$bisimilar, denoted $\mathbb{M}, s \sim \mathbb{M}', s'$ (or $\mathbb{M}, s \sim^n \mathbb{M}', s'$), if every world in s is $(n\text{-})$bisimilar to some world in s' and vice versa. Two models $\mathbb{M}$ and $\mathbb{M}'$ are *globally bisimilar*, denoted $\mathbb{M} \sim \mathbb{M}'$, if every world in $\mathbb{M}$ is bisimilar to some world in $\mathbb{M}'$ and vice versa.

2.4 An inquisitive modal Ehrenfeucht–Fraïssé theorem

In line with the general situation for back-and-forth games that are modelled on the expressive power of given logics, the finite levels of inquisitive bisimulation equivalence capture levels of indistinguishability in terms of inquisitive modal logic. This is brought out in the corresponding Ehrenfeucht–Fraïssé theorem: for finite sets $\mathcal{P}$ of atomic propositions and of agents $\mathcal{A}$, n-bisimilarity $\sim^n$ coincides with $\equiv_{\text{INQML}}^n$, which is logical indistinguishability in INQML at nesting depth up to n, INQML_n. As an immediate consequence, the semantics of all of INQML is preserved under inquisitive bisimulation, which refines all levels $\sim^n$ (in fact, it even strictly refines their common refinement).

Theorem 2.9 (Ehrenfeucht–Fraíssé theorem for INQML [10])**.**
Over any finite sets of agents $\mathcal{A}$ and atomic propositions $\mathcal{P}$, and for any $n \in \mathbb{N}$ and inquisitive epistemic world- or state-pointed models $\mathbb{M}, w$ and $\mathbb{M}', w'$ or $\mathbb{M}, s$ and $\mathbb{M}', s'$:

(i) $\mathbb{M}, w \sim^n \mathbb{M}', w' \iff \mathbb{M}, w \equiv_{\text{INQML}}^n \mathbb{M}', w'$

(ii) $\mathbb{M}, s \sim^n \mathbb{M}', s' \iff \mathbb{M}, s \equiv_{\text{INQML}}^n \mathbb{M}', s'$

The implications from left to right go through (by simple syntactic induction) regardless of the finiteness requirement for $\mathcal{A}$ and $\mathcal{P}$. The finiteness requirement is necessary, though, for the implications from right to left, which can be based on the provision of characteristic formulae $\chi_{\mathbb{M}, w}^n$ and $\chi_{\mathbb{M}, s}^n \in \text{INQML}_n$ that define the $\sim^n$-equivalence class of $\mathbb{M}, w$, and the persistent closure of the $\sim^n$-equivalence class of $\mathbb{M}, s$, respectively;

$$\mathbb{M}', w' \models \chi_{\mathbb{M}, w}^n \iff \mathbb{M}', w' \sim^n \mathbb{M}, w$$
$$\mathbb{M}', s' \models \chi_{\mathbb{M}, s}^n \iff \mathbb{M}', s' \sim^n \mathbb{M}, t \text{ for some } t \subseteq s$$

For the details of the proof, the reader is referred to [10].

Making use of the characteristic formulae, it is easy to obtain the following characterisation of the properties definable in InqML.

Corollary 2.10. *Over a finite signature (i.e. with finite sets of agents $\mathcal{A}$, and of propositions $\mathcal{P}$), a property of world-pointed models is definable in* InqML *if and only if it is closed under $\sim^n$ for some $n \in \mathbb{N}$. Similarly for an inquisitive property of state-pointed models.*

2.5 Relational inquisitive models

The main purpose of this paper lies in the characterisation of inquisitive epistemic logic as a fragment of first-order logic. But the relevant inquisitive models are not directly conceived as relational structures (in the manner that e.g. ordinary Kripke structures can directly be thought of as relational structures for first-order logic). In order to translate the inquisitive setting, which involves a universe of information states over and above the universe of worlds, into a relational structure in which first-order logic similarly gains access to both kinds of entities, we adopt the two-sorted perspective developed in [10]. In this view, W and $\wp(W)$ contribute domains for two distinct sorts.

Definition 2.11. [relational models] For a signature with sets $\mathcal{A}$ of agents and $\mathcal{P} = \{p_i : i \in I\}$ of atomic propositions, a *relational inquisitive epistemic model* is a two-sorted relational structure

$$\mathfrak{M} = (W, S, (E_a)_{a \in \mathcal{A}}, \varepsilon, (P_i)_{i \in I})$$

with disjoint, non-empty universes W, S for the two sorts, with binary relations $E_a, \varepsilon \subseteq W \times S$, and, for $i \in I$, unary $P_i \subseteq W$, satisfying certain conditions listed below. To state these conditions, we introduce the following notation: $E_a[w] := \{s \in S : (w,s) \in E_a\}$ for the set of E_a-successors of w; $\underline{s} := \{w \in W : (w,s) \in \varepsilon\} \subseteq W$ for the set of ε-predecessors of $s \in S$; and $R_a[w] := \bigcup\{\underline{s} : s \in E_a[w]\}$ for the union of these sets across $E_a[w]$ (or, in other words, $R_a := E_a \circ \varepsilon^{-1}$). We can then state the relevant conditions as follows:

  (i) extensionality: if $\underline{s} = \underline{s}'$, then $s = s'$.

  (ii) local powerset: if $s \in S$ and $t \subseteq \underline{s}$, there is an $s' \in S$ such that $\underline{s}' = t$.

and for all $a \in \mathcal{A}$ and $w, w' \in W$:

 (iii) non-emptiness: $E_a[w] \neq \emptyset$.

 (iv) downward closure: if $s \in E_a[w]$, then $t \in E_a[w]$ for all $t \in S$ with $\underline{t} \subseteq \underline{s}$.

 (v) factivity: $w \in R_a[w]$.

 (vi) full introspection: if $w' \in R_a[w]$, then $E_a[w'] = E_a[w]$.

By extensionality, the second sort S of such a relational model can always be identified with a domain of sets over the first sort, namely, $\{\underline{s} : s \in S\} \subseteq \wp(W)$. We always assume this identification and view a relational model as a structure $\mathfrak{M} = (W, S, (E_a), \in, (P_i))$ where $S \subseteq \wp(W)$ and $\in$ is the actual

membership relation; this allows us to specify relational models as just $\mathfrak{M} = (W, S, (E_a), (P_i))$.

With a relational model $\mathfrak{M}$ we associate the Kripke structure $\mathfrak{K}(\mathfrak{M}) = (W, (R_a), (P_i)_{i \in I})$, where R_a is the accessibility relation defined by $R_a[w] = \bigcup_{a \in \mathcal{A}} E_a[w]$ (or, equivalently, by $R_a = E_a \circ \ni$). Note that by virtue of (v) and (vi), each R_a is an equivalence relation on W, and so $\mathfrak{K}(\mathfrak{M})$ is a (multi-agent) $S5$ model. The R_a-equivalence class of w is also denoted as $[w]_a \in W/R_a$.

Further natural constraints on a relational model $\mathfrak{M}$ impose richness conditions on the second sort $S \subseteq \wp(W)$. A relational model $\mathfrak{M} = (W, S, (E_a), (P_i))$ is said to be

- *full* if $S = \wp(W)$;

- *locally full* if for all $a \in \mathcal{A}$ and $w \in W$, $\wp([w]_a) \subseteq S$.

Any relational model $\mathfrak{M} = (W, S, (E_a), (P_i))$ uniquely determines an inquisitive epistemic model $\mathfrak{M}^* = (W, (\Sigma_a), V)$ where $\Sigma_a(w) = E_a[w]$ and $V(p_i) = P_i$. This natural passage from relational models to ordinary models induces semantics for INQML over relational models and supports related natural notions like bisimulation equivalence over these. But as this passage obliterates the explicit (extensional) account of the second sort S, different relational models $\mathfrak{M}$ can determine the same inquisitive model $\mathbb{M}$. In other words, a given inquisitive epistemic model may come with vastly different relational counterparts or *relational encodings*. Here a *relational encoding* of an inquisitive epistemic model $\mathbb{M}$ is any relational model $\mathfrak{M}$ such that $\mathfrak{M}^* = \mathbb{M}$.

For a given inquisitive epistemic model $\mathbb{M}$, the following three relational encodings are particularly natural choices, with minimal or maximal extension for the second sort $S \subseteq \wp(W)$, and – particularly useful for our purposes – an intermediate version with minimal extension among locally full encodings.

Definition 2.12. [relational encodings] For an inquisitive epistemic model $\mathbb{M} = (W, (\Sigma_a), V)$, we define the following relational encodings of $\mathbb{M}$, each based on W as the first sort, and always with $w E_a s \Leftrightarrow s \in \Sigma_a(w)$, $w \, \varepsilon \, s \Leftrightarrow w \in s$ and $P_i = V(p_i)$:

- $\mathfrak{M}^{\mathsf{rel}}(\mathbb{M})$ (minimal) with $S := \bigcup_{a \in \mathcal{A}} \mathrm{image}(\Sigma_a)$;

- $\mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$ (minimal locally full) with $S := \bigcup_{a \in \mathcal{A}} \bigcup_{w \in W} \wp([w]_a)$;

- $\mathfrak{M}^{\mathsf{full}}(\mathbb{M})$ (full, maximal) with $S := \wp(W)$.

Corresponding encodings of state-pointed models $(\mathbb{M}, s)$ augment the respective S by $\wp(s)$.

To illustrate the difference among the three encodings, imagine that a model $\mathbb{M}$ has universe $W = \{w_1, w_2, w_3\}$ and suppose that for each a and each i we have $\Sigma_a(w) = \{\{w_1\}, \{w_2\}, \emptyset\}$. Then

- in the minimal relational encoding $\mathfrak{M}^{\mathsf{rel}}(\mathbb{M})$, $S = \{\{w_1\}, \{w_2\}, \emptyset\}$;

- in the minimal locally full encoding $\mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$, $S = \wp(\{w_1, w_2\})$;

- in the full encoding $\mathfrak{M}^{\mathsf{full}}(\mathbb{M})$, $S = \wp(\{w_1, w_2, w_3\})$.

Not surprisingly, the semantics of INQML over world- or state-pointed inquisitive epistemic models translates neatly into first-order logic over their two-sorted relational encodings, uniformly and regardless of the specific choices for the second sort. A systematic standard translation, as specified in the general inquisitive modal setting in [10], follows the defining clauses in Definition 2.2, reformulated in terms of first-order quantification over the first or second sort of a relational encoding, respectively.[5] Denoting this *standard translation* by $\varphi \mapsto \varphi_{\mathtt{w}}^{\star}$ (with first-order variable symbol $\mathtt{w}$ for worlds, over the first sort) or $\varphi \mapsto \varphi_{\mathtt{s}}^{\star}$ (with first-order variable symbol $\mathtt{s}$ for information states, over the second sort), we find the following.

Observation 2.13. *For any inquisitive epistemic model $\mathbb{M}$, any relational encoding $\mathfrak{M}$ of $\mathbb{M}$, the first-order standard translations $\varphi_{\mathtt{w}}^{\star}$ and $\varphi_{\mathtt{s}}^{\star}$ of $\varphi \in$ INQML are such that for all worlds $w \in W$ and all states $s \in S$:*

(i) $\mathbb{M}, w \models \varphi \iff \mathfrak{M} \models \varphi_{\mathtt{w}}^{\star}[w]$

(ii) $\mathbb{M}, s \models \varphi \iff \mathfrak{M} \models \varphi_{\mathtt{s}}^{\star}[s]$

Modulo this natural standard translation, INQML can be viewed as a syntactic fragment of first-order logic,

$$\text{INQML} \subseteq \mathsf{FO},$$

over the class of all relational inquisitive models. This situation is similar to what the classical standard translation achieves for modal logic ML over Kripke structures, which are naturally conceived as relational structures. The crucial difference for the inquisitive setting is that the target class of structures, unlike the class of relational Kripke structures, forms a non-elementary class. The non-elementary nature of any natural class of relational inquisitive encodings stems form the downward closure condition for inquisitive assignments. This requires representation of the full power set of each information state in the image of Σ in the second sort. That this constraint cannot be captured by FO conditions, can technically be demonstrated by means of compactness arguments for FO. We refer to [10] for related examples, and to [28] for a model-theoretic account that relaxes the notion of (relational) models via passage to a wider class of (relational) pseudo-models, which is elementary and, in a precise sense, conservative w.r.t. INQML semantics.

As a fragment of first-order logic over relational models, i.e. over relational encodings of the intended models, INQML $\subseteq \mathsf{FO}$ displays crucial model-theoretic preservation properties, viz. $\sim$-invariance and, in the state-pointed case, persistency and the empty state property. In showing that INQML $\subseteq \mathsf{FO}$ is *expressively complete* for all those FO-properties of worlds or states (in relational

---

[5]There is a small subtlety here. In [10], two translations are introduced. The first directly mirrors the clauses of the semantics, but in general it only works on encodings which are locally fully. The second, $\varphi \mapsto \varphi^{\star}$, works on arbitrary models, but it requires a pre-processing of the formula φ into an equivalent $\square$-free formula. We refer to [10] for the details.

inquisitive epistemic models from suitable background classes $\mathcal{C}$) that satisfy these semantic constraints, we obtain the desired *characterisation* result. In other words, we want to show that if a FO-formula $\varphi(\mathbf{w})$ in a single free variable of the first sort defines a property of worlds which is bisimulation invariant over $\mathcal{C}$, then it is logically equivalent over $\mathcal{C}$ to some formula of InqML (in defining a world property); and similarly, that if a FO-formula $\varphi(\mathbf{s})$ in a single free variable of the second sort defines a property of information states which, over $\mathcal{C}$, is bisimulation invariant and inquisitive (in the sense of Definition 2.7), then it is logically equivalent over $\mathcal{C}$ to some formula of InqML (in defining a state property). In a nutshell, we seek to establish the semantic equality

$$\text{InqML} \equiv \text{FO}/\!\sim \quad (\dagger)$$

over suitable classes $\mathcal{C}$ of relational epistemic models. Each of these equivalences in expressiveness comes in two readings, one for world properties and one for state properties. We shall mostly in the first instance emphasise the simpler world-pointed reading and discuss the state-pointed versions second.

Such expressive completeness and characterisation results, for world properties in particular, match van Benthem's classical characterisation of modal logic as the bisimulation-invariant fragment of first-order logic. Versions of this characterisation for basic inquisitive modal logic have been expounded in [10]; those results establish the desired expressive completeness results over the classes of (i) all relational inquisitive models, (ii) all finite relational inquisitive models (iii) all locally full models, and (iv) all finite locally full models.[6] A similar expressive completeness result over the class of all full relational inquisitive models can be ruled out by a compactness argument [10]. Due to the non-elementary nature of the relevant classes of two-sorted relational models, the positive results of [10] already need to employ non-classical model-theoretic arguments. Our present analysis for the epistemic case, however, requires a more complex route to expressive completeness, due to the additional constraints imposed on inquisitive epistemic frames, viz. factivity and full introspection.

In preparation for those arguments, we review some salient technical background on the model-theoretic analysis of first-order expressiveness over the kind of two-sorted structures involved here. Important tools in this are Ehrenfeucht-Fraïssé techniques for first-order logic FO as well as for monadic second-order logic MSO.

2.6   Classical Ehrenfeucht–Fraïssé techniques

We review the classical Ehrenfeucht–Fraïssé analysis of the expressive power of FO over relational models. As our relational models are two-sorted, with a second sort $S \subseteq \wp(W)$ that represents subsets of the first sort W, we shall have occasion to look beyond plain first-order expressiveness and to consider also limited forms of monadic second-order expressiveness. The limitations that

---

[6]Relational inquisitive models are defined as in Definition 2.12, but without the requirements of factivity and introspection.

account for the tame nature of MSO applications in our analysis stem from their effective restriction to the local subuniverses $[w]_a$. We note that, for locally full relational models, first-order quantification over the second sort can indeed emulate full MSO in restriction to $[w]_a \subseteq W$, since for these models $\wp([w]_a) \subseteq S$.

Ehrenfeucht–Fraïssé for FO. We review the classical Ehrenfeucht–Fraïssé analysis of the expressive power of FO over relational structures. The underlying game, like the bisimulation game, is a two-player game, that serves to probe, in this case, levels of equivalence w.r.t. first-order logic for matched tuples of elements over these two structures. Positions in the game over relational structures $\mathfrak{A}, \mathfrak{A}'$ are pairs $(\mathbf{a}; \mathbf{a}')$ with marked tuples $\mathbf{a}$ of elements in $\mathfrak{A}$ and $\mathbf{a}'$ of elements in $\mathfrak{A}'$ of the same finite length; we may think of $\mathbf{a}$ and $\mathbf{a}'$ as tuples of elements that are pebbled with pairs of matching pebbles so as to match the components of $\mathbf{a}$ to the components of $\mathbf{a}'$. In a single round, player **I** positions one pebble of the next pair of pebbles in either $\mathfrak{A}$ or $\mathfrak{A}'$, and player **II** must position its match in the opposite structure. This takes the game from position $(\mathbf{a}; \mathbf{a}')$ to some successor position $(\mathbf{a}, a; \mathbf{a}', a')$. Player **II** loses in position $(\mathbf{a}; \mathbf{a}')$ if $\mathfrak{A} \restriction \mathbf{a}, \mathbf{a} \not\simeq \mathfrak{A}' \restriction \mathbf{a}', \mathbf{a}'$, i.e. if the pebbled match does not define an isomorphism between the induced substructures over these elements. If no such position is reached by the end of the game, **II** wins. A winning strategy for **II** in the n-round game starting from position $(\mathbf{a}; \mathbf{a}')$ over $\mathfrak{A}$ and $\mathfrak{A}'$ guarantees that $\mathfrak{A}, \mathbf{a} \equiv_n \mathfrak{A}', \mathbf{a}'$ (n-elementary equivalence), which means that $\mathbf{a}$ in $\mathfrak{A}$ and $\mathbf{a}'$ in $\mathfrak{A}'$ are indistinguishable by FO-formulae of quantifier rank up to n. Denoting the equivalence relation defined in terms of a winning strategy for **II** in the n-round game by $\simeq_n$:

$$\mathfrak{A}, \mathbf{a} \simeq_n \mathfrak{A}', \mathbf{a}' \quad \Rightarrow \quad \mathfrak{A}, \mathbf{a} \equiv_n \mathfrak{A}', \mathbf{a}' \tag{1}$$

For structures $\mathfrak{A}$ and $\mathfrak{A}'$ in a finite relational signature, the classical Ehrenfeucht–Fraïssé theorem states that the converse implication also holds, so that (1) becomes an equivalence, which characterises levels of elementary equivalence in terms of the associated back-and-forth game.

The precise correspondence in the Ehrenfeucht–Fraïssé theorem witnesses how the positioning of a next pebble pair in a round of the game for FO probes the behaviour of the currently matched tuples w.r.t. a single further $\forall/\exists$-quantification over elements – corresponding to an increase of 1 in terms of quantifier rank. The corresponding game for MSO will additionally allow **I** to probe equivalence w.r.t. $\forall/\exists$-quantification over subsets, and game positions correspondingly record a match between elements *and* subsets. In our specific context we shall exploit parallels beween the FO-game involving both sorts of our relational models and the MSO-game over the first sort.

Ehrenfeucht–Fraïssé for MSO. We consider two ordinary structures $\mathfrak{A}, \mathfrak{A}'$ in the same relational signature, with (single-sorted) universes A, A'. To capture the expressiveness of MSO over these, game positions now consist of pairings $(\mathbf{P}, \mathbf{a}; \mathbf{P}', \mathbf{a}')$ where $\mathbf{a}$ and $\mathbf{a}'$ are tuples of elements of the same finite length from

A and A', and $\mathbf{P}$ and $\mathbf{P}'$ are tuples of subsets of A and A' of the same finite length. The condition that $\mathbf{II}$ must maintain in order not to lose in position $(\mathbf{P}, \mathbf{a}; \mathbf{P}', \mathbf{a}')$ over $\mathfrak{A}, \mathfrak{A}'$ now is

$$(\mathfrak{A}, \mathbf{P}) \upharpoonright \mathbf{a}, \mathbf{a} \simeq (\mathfrak{A}', \mathbf{P}') \upharpoonright \mathbf{a}', \mathbf{a}'.$$

This means that the elements matched by the pebble pairs mark out isomorphic induced substructures in the expansions of $\mathfrak{A}$ by the unary predicates $\mathbf{P}$ and of $\mathfrak{A}'$ by $\mathbf{P}'$. These expansions may be seen as colourings produced in second-order moves in which $\mathbf{P}$ and $\mathbf{P}'$ were matched, which are probed (just as in the first-order FO game) by the match in pebbled elements $\mathbf{a}$ and $\mathbf{a}'$ produced in first-order moves. In each round, $\mathbf{I}$ determines whether to play it in first- or second-order mode, and correspondingly makes a choice of either an element or a subset on the $\mathfrak{A}$-side or on the $\mathfrak{A}'$-side; $\mathbf{II}$ must respond likewise on the opposite side. Depending on the mode of this round, it may overall lead from $(\mathbf{P}, \mathbf{a}; \mathbf{P}', \mathbf{a}')$ to $(\mathbf{P}, \mathbf{a}, a; \mathbf{P}', \mathbf{a}', a')$ (with a new pair of pebbled elements (a, a') for first-order round) or to $(\mathbf{P}, P, \mathbf{a}; \mathbf{P}', P', \mathbf{a}')$ (with a new pair of subsets or a new colour for elements, (P, P') for a second-order round).

A winning strategy for $\mathbf{II}$ for n rounds of this game starting from position $(\mathbf{P}, \mathbf{a}; \mathbf{P}', \mathbf{a}')$ over $\mathfrak{A}$ and $\mathfrak{A}'$ now guarantees that $\mathfrak{A}, \mathbf{P}, \mathbf{a} \equiv_n^{\mathsf{MSO}} \mathfrak{A}', \mathbf{P}', \mathbf{a}'$, i.e. indistinguishability by MSO-formulae $\varphi(\mathbf{X}, \mathbf{x})$ of quantifier rank up to n.[7] Denoting the equivalence relation defined in terms of a winning strategy for $\mathbf{II}$ in this n-round game for MSO by $\simeq_n^{\mathsf{MSO}}$, the general result analogous to (1) above is

$$\mathfrak{M}, \mathbf{m} \simeq_n^{\mathsf{MSO}} \mathfrak{M}', \mathbf{m}' \quad \Rightarrow \quad \mathfrak{M}, \mathbf{m} \equiv_n^{\mathsf{MSO}} \mathfrak{M}', \mathbf{m}' \tag{2}$$

and, again, the converse implication is also true over finite relational signatures.

3 Expressive completeness for epistemic InqML

3.1 Bisimulation invariance as a semantic constraint

After the introduction of all the relevant notions, our main theorem announced in Section 1 can be stated more precisely as follows.

THEOREM 1.2. *Let $\mathcal{C}$ be any one of the following classes of relational models: the class of all inquisitive epistemic models; of all finite inquisitive epistemic models; of all locally full, or of all finite locally full inquisitive epistemic models. Over each of these classes,*

$$\textsc{InqML} \equiv \mathsf{FO}/\sim \qquad (\dagger)$$

i.e., a property of world-pointed models is definable in INQML *over $\mathcal{C}$ if and only if it is both* FO*-definable over $\mathcal{C}$ and $\sim$-invariant over $\mathcal{C}$.[8] Similarly, an*

---

[7]We here appeal to a 'mixed' quantifier rank for MSO that does not discriminate between first- and second-order quantification.

[8]As usual, FO-definability of a property $\mathcal{P}$ of world-pointed models over a class $\mathcal{C}$ of models means that there is a FO-formula φ with a single free variable of the first sort such that for all $\mathfrak{M} \in \mathcal{C}$: $(\mathfrak{M}, w) \in \mathcal{P}$ iff $\mathfrak{M} \models \varphi(w)$; similarly for INQML-definability over $\mathcal{C}$ and for properties of state-pointed models and formulae with a free variable of the second sort.

inquisitive property of state-pointed models is definable in INQML *over* $\mathcal{C}$ *if and only if it is* FO*-definable over* $\mathcal{C}$ *and* $\sim$*-invariant over* $\mathcal{C}$.

The inquisitive Ehrenfeucht–Fraïssé theorem, Theorem 2.9, implies that INQML and its standard translation into FO are invariant under bisimulation: INQML $\subseteq$ FO$/\sim$. By Corollary 2.10 it further implies *expressive completeness* of INQML$_n$ for any $\sim^n$-invariant property of world-pointed models and any inquisitive $\sim^n$-invariant property of state-pointed models: FO$/\sim^n \subseteq$ INQML for every $n \in \mathbb{N}$. In order to prove (†) in restriction to some particular class $\mathcal{C}$ of relational inquisitive models in the world-pointed case, it therefore remains to show that, for any formula $\varphi(\mathtt{w}) \in$ FO in a single free variable of the first sort, there is some $n \in \mathbb{N}$ such that $\sim$-invariance of $\varphi(\mathtt{w})$ over $\mathcal{C}$ implies $\sim^n$-invariance over $\mathcal{C}$.

$$\varphi(\mathtt{w}) \in \mathsf{FO}/\sim \implies \varphi(\mathtt{w}) \in \mathsf{FO}/\sim^n \text{ for some } n = n(\varphi) \qquad (\ddagger)$$

This may be viewed as a non-trivial *compactness principle* for $\sim$-invariance of first-order properties, over the non-elementary classes $\mathcal{C}$ of interest.

Observation 3.1. *(a) For any class $\mathcal{C}$ of relational inquisitive models, the following are equivalent:*
    (i) INQML $\equiv$ FO$/\sim$ *for world properties over $\mathcal{C}$;*
    (ii) for FO*-properties of world-pointed models, $\sim$-invariance over $\mathcal{C}$ implies $\sim^n$-invariance over $\mathcal{C}$ for some n.*

(b) Similarly, the following are equivalent:
    (i) INQML $\equiv$ FO$/\sim$ *for inquisitive state properties over $\mathcal{C}$;*
    (ii) for inquisitive FO*-properties of state-pointed models, $\sim$-invariance over $\mathcal{C}$ implies $\sim^n$-invariance over $\mathcal{C}$ for some n.*

As noted above, examples in [10] illustrate that FO does not satisfy a compactness theorem over the non-elementary class of relational inquisitive models. Moreover, over the class of *full* inquisitive relational models, failures of compactness can be exhibited even for the bisimulation-invariant fragment of FO, which must therefore be a strict extension of INQML over this class. In other words, an analogue of Theorem 1.2 fails for the class of full models.

To establish (†) over some class $\mathcal{C}$ of relational models, we aim to show condition (‡) by means of an upgrading of $\sim^n$-equivalence to $\equiv_q$, where q is the quantifier rank of the given formula $\varphi \in$ FO and $n = n(q)$ is a suitable natural number depending on q. If suitable bisimilar partner structures like $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$ in Figure 1 are available within $\mathcal{C}$ for any pair of $\sim^n$-equivalent pointed structures $\mathfrak{M}$ and $\mathfrak{M}'$ from $\mathcal{C}$, then the detour through the lower rung in these diagrams shows the semantics of φ to be $\sim^n$-invariant over $\mathcal{C}$, hence equivalently expressible in INQML over $\mathcal{C}$ due to Observation 3.1.

This general approach to expressive completeness has been tested in several cognate scenarios that range from a simple and constructive proof in [25] of van Benthem's characterisation of ML [36] and Rosen's finite model theory version [33] to the characterisation of INQML as the $\sim$-invariant fragment of FO
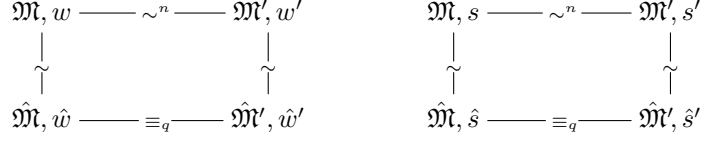
$$\mathfrak{M}, w \ \overline{\quad \sim^n \quad} \ \mathfrak{M}', w' \qquad\qquad \mathfrak{M}, s \ \overline{\quad \sim^n \quad} \ \mathfrak{M}', s'$$

$$\wr \qquad\qquad\qquad \wr \qquad\qquad\qquad\qquad \wr \qquad\qquad\qquad \wr$$

$$\hat{\mathfrak{M}}, \hat{w} \ \overline{\quad \equiv_q \quad} \ \hat{\mathfrak{M}}', \hat{w}' \qquad\qquad \hat{\mathfrak{M}}, \hat{s} \ \overline{\quad \equiv_q \quad} \ \hat{\mathfrak{M}}', \hat{s}'$$

Figure 1: Generic upgrading patterns.

$$\mathfrak{M}, w \ \overline{\quad\quad \sim^n \quad\quad} \ \mathfrak{M}', w' \qquad\qquad \mathfrak{M}, s \ \overline{\quad\quad \sim^n \quad\quad} \ \mathfrak{M}', s'$$

$$\wr \qquad\qquad\qquad\qquad \wr \qquad\qquad\qquad\qquad\qquad \wr \qquad\qquad\qquad\qquad \wr$$

$$\hat{\mathfrak{M}}, \hat{w} \ \overline{\quad\quad \equiv_q \quad\quad} \ \hat{\mathfrak{M}}', \hat{w}' \qquad\qquad \hat{\mathfrak{M}}, \hat{s} \ \overline{\quad\quad \equiv_q \quad\quad} \ \hat{\mathfrak{M}}', \hat{s}'$$

$$\mathsf{I}(\cdot) \downarrow \qquad\qquad\qquad \uparrow \mathsf{I}(\cdot) \qquad\qquad\qquad \mathsf{I}(\cdot) \uparrow \qquad\qquad\qquad \uparrow \mathsf{I}(\cdot)$$

$$\mathsf{I}(\hat{\mathfrak{M}}, \hat{w}) \ \overline{\quad \equiv_{q+d} \quad} \ \mathsf{I}(\hat{\mathfrak{M}}', \hat{w}') \qquad\qquad \mathsf{I}(\hat{\mathfrak{M}}, \hat{s}) \ \overline{\quad \equiv_{q+d} \quad} \ \mathsf{I}(\hat{\mathfrak{M}}', \hat{s}')$$
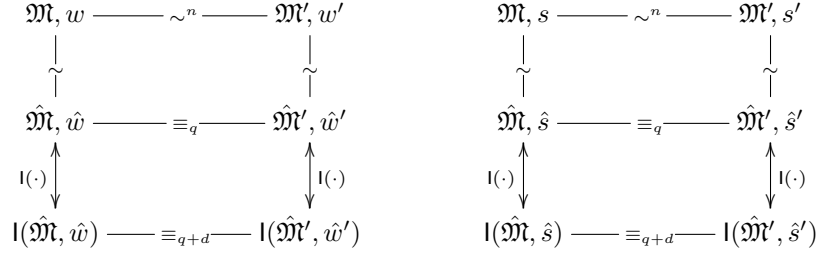
Figure 2: Upgrading patterns for relational epistemic models.

over (finite or general, possibly just locally full) relational inquisitive models in the basic rather than the epistemic setting in [10].

The upgrading argument needs to apply within $\mathcal{C}$, where $\sim$-invariance can be assumed for the vertical transfers. At the same time it must overcome, with the choice of $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$, obstacles to the desired FO-indistinguishability that are not governed by n-bisimilarity. Among those features are issues related to distinctions in low-level multiplicities or local cyclic patterns that can be expressed in FO at quantifier rank q. So within the classes $\mathcal{C}$ of relational epistemic models that are of interest here, we need to come up with bisimilar partner structures $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$, for which Ehrenfeucht–Fraïssé arguments can establish $\equiv_q$. That task can be further split into two main stages:

(1) the construction of suitable bisimilar inquisitive epistemic companion models $\hat{\mathbb{M}} \sim \mathfrak{M}$ and $\hat{\mathbb{M}}' \sim \mathfrak{M}'$, and

(2) a model-theoretic interpretation of the relevant relational encodings $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$ of $\hat{\mathbb{M}}$ and $\hat{\mathbb{M}}'$ in related two-sorted first-order structures $\mathsf{I}(\hat{\mathbb{M}}) = \mathsf{I}(\hat{\mathfrak{M}})$ and $\mathsf{I}(\hat{\mathbb{M}}') = \mathsf{I}(\hat{\mathfrak{M}}')$; this yields a neater argument for $\equiv_q$ involving, among other techniques, locality arguments based on Gaifman's theorem for FO.

We first review and prepare the ground for (2) in Section 3.2 below; then address (1) in Section 3.3; and finally combine these preparatory steps to prove our main characterisation theorem, Theorem 1.2, in Section 3.4. The picture for the intended upgrading arguments can be extended as shown in Figure 2.

3.2 First-order expressiveness over relational models

As pointed out above, the assessment of the expressive power of FO over relevant non-elementary classes of relational structures cannot use classical compactness arguments. Instead it draws on Ehrenfeucht–Fraïssé arguments. These rely on *locality criteria* based on *Gaifman distance* (cf. Definition 3.6) and suitable local tree-like unfoldings in appropriate image structures $I(\mathfrak{M})$ – structural correlates that lend themselves to the localisation of Ehrenfeucht–Fraïssé arguments, as it were, and related to $\mathfrak{M}$ by mutual FO-interpretability.

First-order interpretations in an exploded view. Consider an epistemic model $\mathbb{M} = (W, (\Sigma_a)_{a \in \mathcal{A}}, V)$ with locally full two-sorted relational encoding $\mathfrak{M} := \mathfrak{M}^{\mathsf{lf}}(\mathbb{M}) = (W, S, (E_a)_{a \in \mathcal{A}}, (P_i)_{i \in I})$, where

$$S = \bigcup \{\wp([w]_a) \colon w \in W, a \in \mathcal{A}\},$$

where $[w]_a = \sigma_a(w) = R_a[w]$. We also consider the natural restrictions-cum-reducts $\mathbb{M} \upharpoonright [w]_a$ and $\mathfrak{M} \upharpoonright [w]_a$ for any combination of $w \in W$ and $a \in \mathcal{A}$:

$$\mathbb{M} \upharpoonright [w]_a := \big([w]_a, \Sigma_a \upharpoonright [w]_a, V \upharpoonright [w]_a\big)$$

will be referred to as an *a-local structure* or just *local structure* and is treated as inquisitive model for the single agent a with constant Σ_a. The relational encoding

$$\mathfrak{M} \upharpoonright [w]_a := \big([w]_a, \wp([w]_a), E_a \upharpoonright [w]_a, (P_i \upharpoonright [w]_a)\big)$$

of this local structure arises as the induced substructure of the reduct of $\mathfrak{M}$ that eliminates all other agents; this restriction is two-sorted and full, as the domain for the second sort, the restriction of S to $\wp([w]_a)$ comprises all of $\wp([w]_a)$ since $\mathfrak{M} = \mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$ is locally full. So each $\mathfrak{M} \upharpoonright [w]_a$ is a full relational encoding of the single-agent, single-equivalence-class inquisitive epistemic model $\mathbb{M} \upharpoonright [w]_a$. In order to analyse the expressive power of first-order logic over $\mathfrak{M}$, we look at an "exploded view" $I(\mathfrak{M})$ of $\mathfrak{M}$, which represents $\mathfrak{M}$ as the conglomerate of these restrictions $\mathfrak{M} \upharpoonright [w]_a$, with their interconnections charted in a core structure representing the underlying Kripke frame. The concrete format of such an exploded view would admit quite some variations. What matters for us is first-order interpretability of $\mathfrak{M}$ itself within $I(\mathfrak{M})$: as indicated in Figure 2, we want to apply these representations to suitably prepared models $\hat{\mathfrak{M}}$ so as to argue for $\equiv_q$ at the level of $\hat{\mathfrak{M}}/\hat{\mathfrak{M}}'$ on the basis of $\equiv_{q+d}$ at the level of the $I(\hat{\mathfrak{M}})/I(\hat{\mathfrak{M}}')$ for some fixed constant d. We first discuss suitable representations $I(\mathfrak{M})$ for arbitrary locally full relational epistemic models of the form $\mathfrak{M} = \mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$ but note that they will only be used for specifically prepared bisimilar partner structures as in Figure 2; it is also important to note that the $I(\cdot)$-images are not themselves relational encodings of inquisitive models. Figure 3 illustrates the structural ingredients in the format that we choose for this exploded view $I(\mathfrak{M})$ of $\mathfrak{M}$. The two-sorted structure $I(\mathfrak{M})$, from which $\mathfrak{M}$ itself can be retrieved by first-order means, involves (see below for details):

– a *core part* within the first sort, which represents the underlying multi-agent $S5$ Kripke frame on the set of worlds W, with relation symbols R_a for accessibility w.r.t. agent $a \in \mathcal{A}$ as in the induced Kripke frame;

– a halo of attached disjoint external copies of each one of the two-sorted *local structures* $\mathfrak{M}\upharpoonright[w]_a$, one for each equivalence class of worlds w.r.t. to agents $a \in \mathcal{A}$; $\mathfrak{M}\upharpoonright[w]_a$ carries the natural restrictions of the relations in $\mathfrak{M}$ but with inquisitive assignment just for agent a (see above);

– with additional unary marker predicates P_a for the a-local components and a binary relation R_I for the identification between worlds in the core part and their instances in local component structures in the halo.

The first part of the disjoint union underlying $\mathsf{I}(\mathfrak{M})$ forms the core (fully in the first sort), the remaining parts (all truly two-sorted) form the halo, with the interpretation of R_I as the only link between the otherwise disjoint pieces:

$$\mathsf{I}(\mathfrak{M}) := \big((W, (R_a)_{a \in \mathcal{A}}) \ \oplus \bigoplus_{a \in \mathcal{A}, [w]_a \in W/R_a} \mathfrak{M}\upharpoonright[w]_a \ , (P_a)_{a \in \mathcal{A}}, \ R_I\big)$$

We think of the disjoint union of the restrictions $\mathfrak{M}\upharpoonright[w]_a$ as realised over a-tagged versions of $[w]_a$ with elements u_a for $u \in [w]_a$ for the first sort and $[w]_a$-tagged elements s_α for $s \in \wp([w]_a)$ in the second sort for $\alpha = [w]_a \in W/R_a$. More formally these can be realised over universes $[w_a] \times \{a\}$ and $\wp([w]_a) \times \{\alpha\}$, but we use notation like u_a instead of (u, a) for better readability.[9] The interpretations for E_a and the valuation V are just restricted over $[w]_a$ as indicated above. The core structure $(W, (R_a))$ in the first-sort of $\mathsf{I}(\mathfrak{M})$ is the underlying Kripke frame on the set of worlds W with the modal accessibility relations (representations of the σ_a as equivalence relations)

$$R_a \ := \ \{(w, w') \in W^2 \colon w' \in \sigma_a(w)\}$$

that make up this $S5$ frame; its domain W also serves to supply the markers to relate any $v \in W$ with its multiple copies in the local structures $\mathfrak{M}\upharpoonright[w]_a$ through R_I; the extra unary relation symbols P_a carve out the subset of a-tagged copies w_a of worlds $w \in W$,

$$P_a \ := \ \{u_a \colon u \in [w]_a\},$$
$$R_I \ := \ \{(w, w_a) \colon w \in W, a \in \mathcal{A}\}.$$

One may read the index I for "incidence" or "identification" of worlds with their multiple incarnations; the subset W of the first sort of $\mathsf{I}(\mathfrak{M})$ precisely serves as a set of world labels that relate copies of the corresponding w:

$$u_a \text{ represents the same world as } u'_{a'} \text{ iff } (u_a, u'_{a'}) \in (R_I)^{-1} \circ R_I,$$

i.e. iff there is some $v \in W$ for which $u_a, u'_{a'} \in R_I[v]$, which implies that $v = u = u'$ and that $[w]_a = [v]_a$ and $[w']_{a'} = [v]_{a'}$ for this $v = u = u'$.

---

[9]One could also tag the elements $u_a = (u, a)$ for $u \in [w]_a$ slightly redundantly by their $\alpha = [w]_a \in W/R_a$; similarly, the α-tag of $\emptyset \neq s \in \wp([w]_a)$ in $\mathfrak{M}\upharpoonright[w]_a$ is determined by a, since $u_a \in s$ implies $\alpha = [u]_a = [w]_a$, but we do distinguish $\emptyset \in \wp([w]_a)$ from $\emptyset \in \wp([w']_{a'})$ unless $a = a'$ and $[w']_{a'} = [w]_a \in W/R_a$.
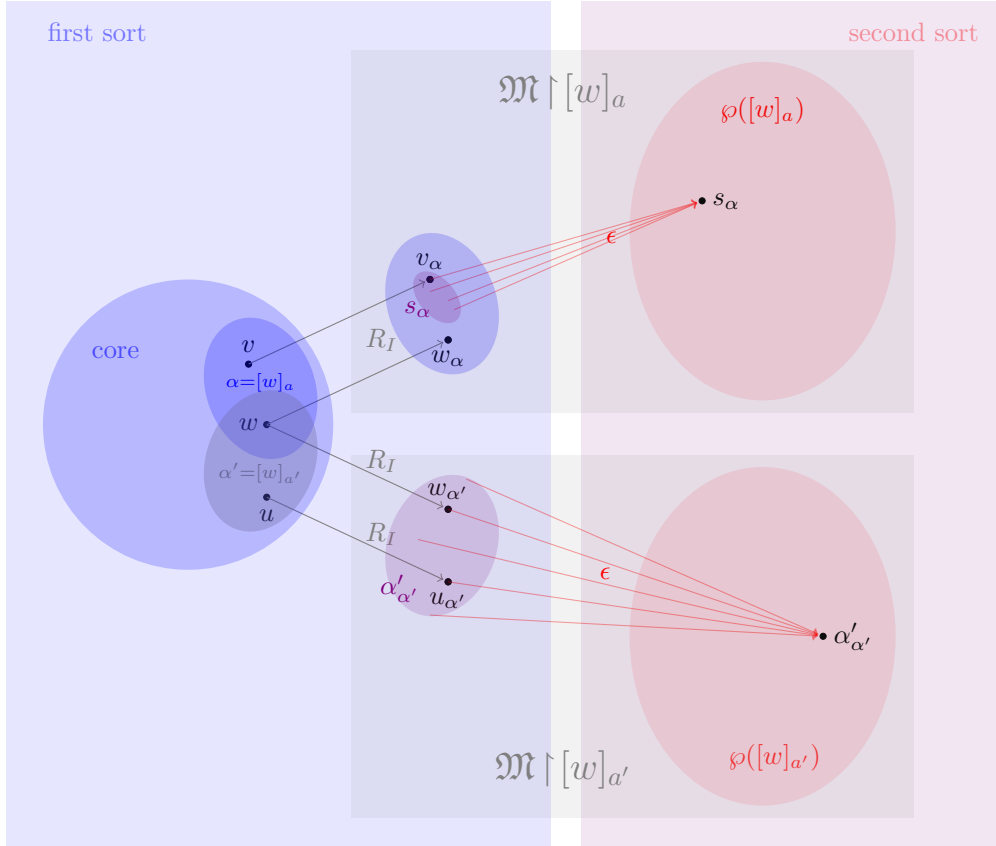
Figure 3: Structural layout of the exploded view $\mathsf{I}(\mathfrak{M})$ of $\mathfrak{M}$; selectively displayed are the representations of intersecting local structures $\mathfrak{M} \restriction [w]_a$, $\mathfrak{M} \restriction [w]_{a'}$ w.r.t. agents a and a', with one element each from their (full powerset) second sorts. Here $\alpha = [w]_a$ and $\alpha' = [w]_{a'}$ contribute two overlapping local structures, whose disjoint (tagged) representations each contribute the full power set of their first sort to disjoint (tagged) representations in the second sort; displayed as particular elements of the second sort are s_α for some $s \subseteq \alpha$ (as part of the representation of the a-local structure $\mathfrak{M} \restriction [w]_a$) and the full set $\alpha' \subseteq \alpha'$ (as part of the representation of the a'-local structure $\mathfrak{M} \restriction [w]_{a'}$).

The world-pointed model $\mathfrak{M}, w$ is naturally represented by the world-pointed exploded view, $\mathsf{I}(\mathfrak{M}), w$ with distinguished world w in the core of $\mathsf{I}(\mathfrak{M})$. In the state-pointed $\mathfrak{M}, s$ with nonempty state $s \in S$, s need not be directly represented as an element of the second sort in any local structure, as in general we only require $\wp(s) \subseteq S$ (cf. Definition 2.12). But we may reduce this scenario to the world-pointed case by the introduction of an extra (dummy) agent as follows.[10] For non-trivial $s \subseteq W$ where $\wp(s) \subseteq S$ let $\mathcal{A}_s := \mathcal{A} \dot\cup \{s\}$ and consider the expansion of $\mathfrak{M}$ to $\mathfrak{M}_s := (\mathfrak{M}, E_s)$ based on the relational encoding of the trivial inquisitive assignment

$$\Sigma_s \colon w \mapsto \left\{ \begin{array}{ll} \wp(s) & \text{for } w \in s \\ \wp(\{w\}) & \text{for } w \notin s \end{array} \right.$$

whose induced σ_s and R_s partition W into singleton equivalence classes $[w]_s = \{w\}$ for $w \notin s$ and the single, potentially non-trivial equivalence class $[w]_s = s$, for any $w \in s$. So the exploded view of $\mathfrak{M}_s$ has, for non-trivial s, an extra part for the s-local structure $\mathfrak{M}_s \restriction [w]_s$ which provides a representation of s, both as a set of worlds in its first sort, and as the $\subseteq$-maximal element in its second sort. So s is FO-definably available in $\mathsf{I}(\mathfrak{M}_s)$ in the s-local structure $\mathfrak{M}_s \restriction [w]_s$, marked out either by the parameter $s = s_{\alpha_s}$, $\alpha_s := [w]_s$, of the second sort, or by any choice of a parameter w_s, $w \in s$, in the second sort.

We note that passage to $\mathfrak{M}_s$ is safe also w.r.t. bisimulation equivalences for state-pointed models in the sense that $\mathfrak{M}, s \sim \mathfrak{M}', s'$ implies that $\mathfrak{M}_s, s \sim \mathfrak{M}'_{s'}, s'$, i.e. that for any $w \in s$ there is some $w' \in s'$ such that $\mathfrak{M}_s, w \sim \mathfrak{M}'_{s'}, w'$ and vice versa; and similarly for every level of $\sim^n$. So we give the following definition.

Definition 3.2. The *exploded view* of a locally full relational epistemic model $\mathfrak{M}$ is defined as $\mathsf{I}(\mathfrak{M})$ according to the above; the analogous representation of an epistemic model $\mathbb{M}$ is obtained as $\mathsf{I}(\mathbb{M}) := \mathsf{I}(\mathfrak{M}^{\mathsf{lf}}(\mathbb{M}))$. For world-pointed models we put $\mathsf{I}(\mathfrak{M}, w) := \mathsf{I}(\mathfrak{M}), w$. In the non-trivially state-pointed version we use $\mathsf{I}(\mathfrak{M}, s) := \mathsf{I}(\mathfrak{M}_s), s$ where $\mathfrak{M}_s$ is the expansion for a dummy agent s, and s stands for its representation s_{α_s} in the second sort of $\mathfrak{M} \restriction \alpha_s$ as discussed above.

The exploded views $\mathsf{I}(\mathfrak{M}, w)$ and $\mathsf{I}(\mathfrak{M}, s)$ are such that the evaluation of any FO-formula in a single free first- or second-sort variable over $\mathfrak{M}, w$ or $\mathfrak{M}, s$ translates into the evaluation of a corresponding first-order formula over $\mathsf{I}(\mathfrak{M}, w)$ and $\mathsf{I}(\mathfrak{M}, s)$, respectively. Indeed, there is a uniform syntactic translation to this effect, a phenomenon that generally applies for first-order interpretations. In our case, we rely on the uniform FO interpretability of $\mathfrak{M}$ within its exploded view $\mathsf{I}(\mathfrak{M})$, together with the FO-definable access to the relevant assignment w or s: $w \in W$ is an element of the first sort of $\mathsf{I}(\mathfrak{M}, w)$; and $s \neq \emptyset$ is represented by the maximal element in the second sort of its local component structure.[11]

---

[10] The case of $\emptyset$-state-pointed models is trivial, since the empty information state is already trivially (and multiply) represented in the component structures $\mathfrak{M} \restriction [w]_a$; so are all singleton information states.

[11] A standard technical detail involves the passage to quotients w.r.t. FO-definable equiv-

In fact, for locally full inquisitive relational models $\mathfrak{M}$, the two-sorted structures $\mathfrak{M}$ and their exploded views $\mathsf{I}(\mathfrak{M})$ are uniformly first-order interpretable within one another (bi-interpretable). The following summarises the use we want to make of interpretability in the context of the upgrading scheme as indicated in Figure 2.

Claim 3.3. *Due to FO-interpretability of any locally full relational model $\mathfrak{M}, w$ in $\mathsf{I}(\mathfrak{M}, w)$, or of $\mathfrak{M}, s$ in $\mathsf{I}(\mathfrak{M}, s)$, there is a constant d such that $(q + d)$-elementary equivalence between the exploded views of two such models implies q-elementary equivalence between the original relational models. For instance in the world-pointed case, any first-order $\varphi(\mathtt{w})$ of quantifier rank up to q in a world-variable has a uniform translation $\varphi_I(\mathtt{w})$ of quantifier rank at most $q + d$ such that $\mathfrak{M}, w \models \varphi$ iff $\mathsf{I}(\mathfrak{M}, w) \models \varphi_I$. It follows that for any pair of world-pointed locally full inquisitive relational models $\mathfrak{M}, w$ and $\mathfrak{M}', w'$*

$$\mathsf{I}(\mathfrak{M}, w) \equiv_{q+d} \mathsf{I}(\mathfrak{M}', w') \quad \Rightarrow \quad \mathfrak{M}, w \equiv_q \mathfrak{M}', w';$$

and similarly for formulae $\varphi(\mathtt{s})$ in a state variable and state-pointed locally full relational epistemic models.

Proofsketch. All the relevant structural information about $\mathfrak{M}$ is easily imported to the sub-universe W of the first sort in $\mathsf{I}(\mathfrak{M})$. This sub-universe itself is FO definable in $\mathsf{I}(\mathfrak{M})$ as the projection of R_I to its first component; the valuation V at $w \in W$ is imported from any one of the elements in $R_I[w]$; the relational encoding $E_a[w]$ of the inquisitive assignment $\Sigma_a(w)$ is similarly imported through backward translation of $E_a(w_a)$ from any one of the $w_a \in R_I[w] \cap P_a$ via set-wise passage to R_I-predecessors:

$$v \in s \in \Sigma_a(w) \quad \text{iff} \quad v \in (R_I \circ \in)^{-1}(E_a[w_a])$$

for $w_a \in R_I[w] \cap P_a$. Compare Figure 3 to see how especially this relationship between elements and subsets of the core part can be traced in the representation of the a-local structure $\mathfrak{M} \restriction [w]_a$ in the halo of $\mathsf{I}(\mathfrak{M})$. All these relations are therefore uniformly first-order definable within the exploded views $\mathsf{I}(\mathfrak{M})$, and the maximal quantifier rank in corresponding FO-formulae can serve as the value d for the transfer claim. For the non-trivially state-pointed case, similarly, the distinguished state s (and its entire powerset) is accessible in the component structure $\mathfrak{M}_s \restriction [w]_s$ for any $w \in s$ in $\mathsf{I}(\mathfrak{M}, s)$. $\qquad \square$

For the upgrading argument according to Figure 2 we shall therefore eventually establish $\equiv_q$-equivalence of, for instance world-pointed inquisitive relational models,

$$\mathfrak{M}, w \equiv_q \mathfrak{M}', w' \tag{3}$$

alence relations for the representation of especially the second sort of $\mathfrak{M}$ with $\mathsf{I}(\mathfrak{M})$. By the very nature of an "exploded view", the empty information state $\emptyset \in S$, all singleton states $\{w\} \in S$ and possibly other information states are represented in the second sort of more than one local structure $\mathfrak{M} \restriction [w]_a$. The corresponding representation of the first sort W is trivial because W is included as a definable subset in the first sort of $\mathsf{I}(\mathfrak{M})$, even though structural properties of its elements need to be retrieved from R_I-related copies in the $\mathfrak{M} \restriction [w]_a$ (see proofsketch below).

where q is the quantifier rank of the given bisimulation invariant $\varphi(x) \in \mathsf{FO}$, based on the assumption that $\mathfrak{M}, w \sim^\ell \mathfrak{M}', w'$ for suitable $\ell = \ell(q)$. The above preparation allows us to reduce this to showing

$$\mathsf{I}(\mathfrak{M}, w) = \mathsf{I}(\mathfrak{M}), w \equiv_{q+d} \mathsf{I}(\mathfrak{M}'), w' = \mathsf{I}(\mathfrak{M}', w'). \tag{4}$$

In fact we also only need to establish (3) or (4) for $\mathfrak{M}$ and $\mathfrak{M}'$ from some subclass of the class of all epistemic relational models that represents the corresponding class (of all or just all finite such models) up to bisimulation equivalence. E.g. in the finite model theory version for the world-pointed case, it suffices to show that, for any two finite epistemic models $\mathbb{M}, w$ and $\mathbb{M}', w'$ that are ℓ-bisimilar, there are bisimilar companions $\hat{\mathbb{M}}, \hat{w} \sim \mathbb{M}, w$ and $\hat{\mathbb{M}}', \hat{w}' \sim \mathbb{M}', w'$ such that

$$\mathsf{I}(\hat{\mathbb{M}}, \hat{w}) = \mathsf{I}(\mathfrak{M}^{\mathsf{lf}}(\hat{\mathbb{M}})), \hat{w} \equiv_{q+d} \mathsf{I}(\mathfrak{M}^{\mathsf{lf}}(\hat{\mathbb{M}}')), \hat{w}' = \mathsf{I}(\hat{\mathbb{M}}', \hat{w}') \tag{5}$$

The remaining work towards this upgrading argument, therefore, focuses on the construction of suitable bisimilar companions. Suitability here means that (5) can be established by Ehrenfeucht–Fraïssé arguments for FO and – for auxiliary arguments concerning the restrictions $\mathfrak{M} \restriction [w]_a$ as main building blocks of the $\mathsf{I}(\mathfrak{M})$ – for MSO.

The Ehrenfeucht–Fraïssé argument for suitably prepared epistemic models is facilitated by the observation that $\sim$-invariance implies invariance under disjoint unions for epistemic models, and that the passage to exploded views is compatible with such disjoint unions. The following definition of disjoint unions is the natural one for inquisitive models.

Definition 3.4. The *disjoint union* of epistemic models $\mathbb{M}_1$, $\mathbb{M}_2$, denoted as $\mathbb{M}_1 \oplus \mathbb{M}_2$, is the epistemic model whose set of worlds is the disjoint union of the sets of worlds W_i of the $\mathbb{M}_i$, with the obvious transfer of inquisitive and propositional assignments from the $\mathbb{M}_i$ to $\mathbb{M}_1 \oplus \mathbb{M}_2$.

Clearly, $\mathbb{M}, w \sim \mathbb{M} \oplus \mathbb{L}, w$ and $\mathbb{M}, s \sim \mathbb{M} \oplus \mathbb{L}, s$ for any world- or state-pointed epistemic models $\mathbb{M}, w$ or $\mathbb{M}, s$ and plain $\mathbb{L}$. But while the disjoint union of epistemic models or of their relational encodings cannot be strictly disjoint, because of the special rôle of the empty information state $\emptyset$, this little problem is overcome in their exploded views.

Observation 3.5. *Passage to the exploded view of epistemic models is compatible with disjoint unions in the sense that*

$$\mathsf{I}(\mathbb{M}_1 \oplus \mathbb{M}_2) = \mathsf{I}(\mathbb{M}_1) \oplus \mathsf{I}(\mathbb{M}_2),$$

and similarly with a distinguished world or state parameter from $\mathbb{M}_1$ say.

FO-equivalence and Gaifman locality. We review some relevant technical background on Gaifman locality and its use towards establishing levels of first-order equivalence between relational structures, with a view to applying

24

these techniques to show $I(\hat{\mathfrak{M}}, \cdot) \equiv_{q+d} I(\hat{\mathfrak{M}}', \cdot)$ in the situation of Figure 2, for suitably pre-processed models $\hat{\mathfrak{M}}, \cdot \sim^n \hat{\mathfrak{M}}', \cdot$. The relational structures in question interpret just unary and binary relations. Correspondingly, *Gaifman distance* over their (two-sorted) universes is just ordinary graph distance w.r.t. the symmetrisations of the binary relations. This establishes a natural distance measure between elements, in our case across both sorts, worlds and information states.[12] For an element b of the relational structure $\mathfrak{B}$ and $\ell \in \mathbb{N}$, the set of elements at distance up to ℓ from b is the *ℓ-neighbourhood* of b, denoted as

$$N^\ell(b) = \{b' \in B \colon d(b, b') \leqslant \ell\}.$$

A first-order formula $\psi(x)$ in a single free variable x (of either sort) is *ℓ-local* if its semantics at $b \in \mathfrak{B}$ depends just on the induced substructure $\mathfrak{B} \restriction N^\ell(b)$ on the ℓ-neighbourhood $N^\ell(b)$. In other words, $\psi(x)$ is ℓ-local if, and only if, it is logically equivalent to its *relativisation* to the ℓ-neighbourhood of x, i.e., to the formula $\psi^\ell(x)$ obtained by explicitly restricting all quantifiers in $\psi(x)$ to the ℓ-neighbourhood of x (which is possible provided the underlying relational signature is finite). We refer to $\psi^\ell(x) \in \mathsf{FO}$ as the *ℓ-localisation* of $\psi(x)$. A finite set of elements of the relational structure $\mathfrak{B}$ is said to be *ℓ-scattered* if the ℓ-neighbourhoods of any two distinct members are disjoint. A *basic local sentence* is a sentence that asserts, for some formula $\psi(x)$ and some $m \geqslant 1$, the existence of an ℓ-scattered set of m elements that each satisfy the ℓ-localisation $\psi^\ell(x)$. Gaifman's theorem [16, 15], in the case that we are interested in, asserts that any first-order formula in a single free variable x is logically equivalent to a boolean combination of ℓ-local formulae $\psi^\ell(x)$ for some ℓ and some basic local sentences. We shall focus on the following levels of Gaifman equivalence.

Definition 3.6. Two pointed relational structures $\mathfrak{B}, b$ and $\mathfrak{B}', b'$ are *ℓ-locally r-equivalent*, denoted $\mathfrak{B}, b \equiv_r^{(\ell)} \mathfrak{B}', b'$, if b and b' satisfy exactly the same ℓ-localisations of formulae $\psi(x)$ of quantifier rank up to r.

$\mathfrak{B}, b$ and $\mathfrak{B}', b'$ are *(ℓ, r, m)-Gaifman-equivalent*, denoted $\mathfrak{B}, b \equiv_{r,m}^{(\ell)} \mathfrak{B}', b'$, if $\mathfrak{B}, b$ and $\mathfrak{B}', b'$ are ℓ-locally r-equivalent and $\mathfrak{B}$ and $\mathfrak{B}'$ satisfy exactly the same basic local sentences concerning ℓ'-localisations for formulae $\psi(x)$ of quantifier rank up to r and ℓ'-scattered sets of sizes up to m for any $\ell' \leqslant \ell$.

The following is an immediate corollary of Gaifman's classical theorem [16, 15], also in the case of our two-sorted structures and for distinguished elements of either sort.

Proposition 3.7. *The semantics of any first-order formula $\psi(x)$ in a finite purely relational vocabulary is preserved under Gaifman equivalence $\equiv_{r,m}^{(\ell)}$ for sufficiently large values of the parameters $\ell, r, m \in \mathbb{N}$. For formulae $\psi(x) \in \mathsf{FO}$ whose semantics is also preserved under disjoint unions of relational structures, preservation under $\equiv_{r,m}^{(\ell)}$ implies preservation under $\equiv_r^{(\ell)}$.*

---

[12]In general, 'infinite distances' occur between elements that are not linked by finite chains of binary relational edges, i.e. that are in distinct connected components.
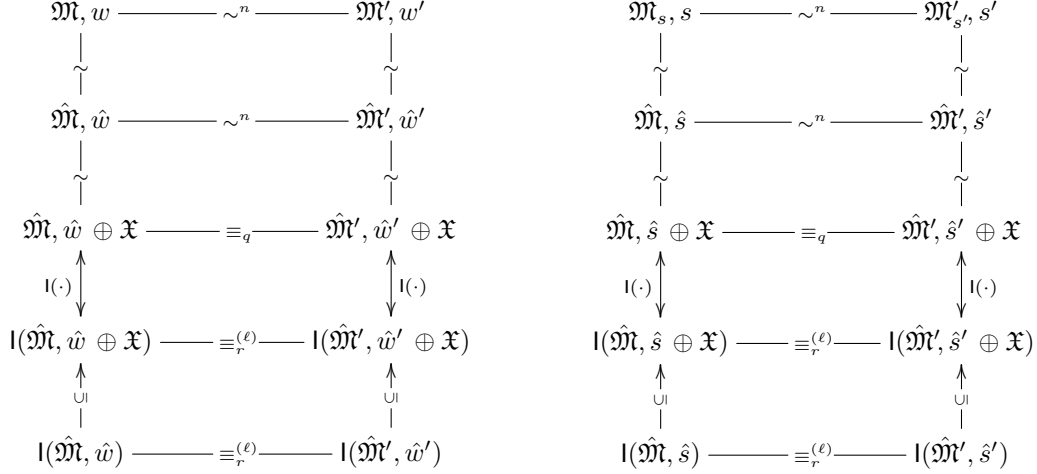
Figure 4: Upgrading for relational epistemic models, refined by Proposition 3.7.

We can use these results from classical first-order model theory in order to simplify the upgrading task: in effect we may replace the target levels $\equiv_q$ from Figure 1, or $\equiv_{q+d}$ in Figure 2 by $\equiv_r^{(\ell)}$ for exploded views, for appropriate ℓ and r. Compare Figure 4 for the following argument. On the basis of interpretability, $\equiv_q$ over $\hat{\mathfrak{M}}/\hat{\mathfrak{M}}'$ follows from $\equiv_{q+d}$ over $\mathsf{I}(\hat{\mathfrak{M}})/\mathsf{I}(\hat{\mathfrak{M}})'$ as illustrated in Figure 2, which in turn is implied by a suitable level of $\equiv_{r,m}^{(\ell)}$ over these. Up to bisimilarity, we may replace models $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$ by disjoint sums with a model

$$\mathfrak{X} \; := \; m \otimes \hat{\mathfrak{M}} \; \oplus \; m \otimes \hat{\mathfrak{M}}'$$

consisting of m many disjoint copies of each of $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$ (and similarly, using $\hat{\mathfrak{M}}_{\hat{s}}$ and $\hat{\mathfrak{M}}'_{\hat{s}'}$ in the non-trivially state-pointed case). The purpose of this modification is simply to trivialise the contribution of basic local sentences about scattered sets of size up to m towards $\equiv_{r,m}^{(\ell)}$. By compatibility of the exploded view with disjoint sums, $\equiv_{r,m}^{(\ell)}$ and $\equiv_r^{(\ell)}$ coincide for the exploded views of these augmented structures, which further reduces the issue to the ℓ-local equivalence

$$\mathsf{I}(\hat{\mathfrak{M}}, \cdot) \; \equiv_r^{(\ell)} \; \mathsf{I}(\hat{\mathfrak{M}}', \cdot) \tag{6}$$

in both the world-pointed and the state-pointed case, as illustrated in Figure 4.

The following summarises these considerations, which pave the way towards establishing the compactness property which we have seen to be equivalent to our main theorem. We can always assume that the underlying signature is the signature of the given first-order formula, hence finite.

Proposition 3.8. *In order to show that any first-order formula in a single free world or state variable that is $\sim$-invariant over a class $\mathcal{C}$ of world- or state-*

pointed relational epistemic models is even invariant under $\sim^n$ for some finite level $n \in \mathbb{N}$, and hence expressible in INQML *over* $\mathcal{C}$, *it suffices to show that, for any $\ell, r \in \mathbb{N}$ there is an $n \in \mathbb{N}$ such that any two $\sim^n$-equivalent models from* $\mathcal{C}$ *can, up to $\sim$, be replaced within* $\mathcal{C}$ *by models, whose exploded views are ℓ-locally* FO-*equivalent up to quantifier rank r.*

Towards the criteria in Proposition 3.8, we exploit the representation of the relevant models in their exploded views. These allow us to view an inquisitive epistemic model $\mathbb{M}$ as a network of interconnected a-local structures $\mathbb{M} \restriction [w]_a$ that are linked through shared worlds. The individual mono-modal a-local structures, where two-sortedness reflects the higher-order nature of the inquisitive assignment, can be managed well in isolation. The interaction between these a-local structures, for different agents a in the global pattern of overlapping a-classes, on the other hand, is essentially reflected in the single-sorted $S5$-Kripke structure $\mathfrak{K}(\mathbb{M})$ that forms the core part of the exploded view $\mathsf{I}(\mathbb{M})$ in the first sort. The desired upgrading argument for Proposition 3.8 calls for a strategy in the r-round first-order Ehrenfeucht–Fraïssé game over the ℓ-neighbourhoods of exploded views of two suitable pre-processed n-bisimilar models. Their nature as conglomerates of interconnected a-local structures allows us to divide concerns strategically as follows:

- a-local concerns in suitably regularised local structures: the analysis of the FO-game over these two-sorted structures largely reduces to MSO-analysis over their first sorts; local pre-processing towards a normalisation of the inquisitive assignments will support that analysis.

- the actual Ehrenfeucht–Fraïssé game in the ℓ-neighbourhoods, involving the connectivity pattern of links between a-local component structures; for that, global pre-processing towards local acyclicity in the connectivity patterns supports an analysis based on tree-like hypergraphs.

And indeed, the structure of our exploded views is precisely designed for this purpose. Due to suitable global pre-processing the core parts will be locally tree-like, and the general format of the exploded view as illustrated in Figure 3 retains this tree-likeness because local structures are connected only through the core; and the tentacles into the second sort, which each belong to just one local structure, will have been tamed by local pre-processing. Both the local and the global pre-processing need to respect bisimulation equivalence, i.e. rely on the construction of suitable bisimilar companions within the class $\mathcal{C}$ at hand.

3.3   Suitable bisimilar companions

To deal with various uniform constructions of bisimilar companions we use the notion of (globally bisimilar) coverings as expounded e.g. in [26, 14, 27]. The natural adaptation to the setting of multi-modal inquisitive epistemic models is the following.

Definition 3.9. A *bisimilar covering* of an inquisitive model $\mathbb{M} = (W, (\Sigma_a), V)$ by an inquisitive model $\hat{\mathbb{M}} = (\hat{W}, (\hat{\Sigma}_a), \hat{V})$ is a map

$$\pi \colon \hat{\mathbb{M}} \longrightarrow \mathbb{M},$$

based on a surjection $\pi \colon \hat{W} \to W$ with natural induced maps $\pi \colon \wp(\hat{W}) \to \wp(W)$ and $\pi \colon \wp\wp(\hat{W}) \to \wp\wp(W)$, such that π is compatible with the inquisitive and propositional assignments in the sense that the following diagram commutes

$$
\begin{array}{ccc}
\hat{W} & \xrightarrow{\;\hat{\Sigma}_a\;} & \wp\wp(\hat{W}) \\
\big\downarrow{\scriptstyle \pi} & & \big\downarrow{\scriptstyle \pi} \\
W & \xrightarrow{\;\Sigma_a\;} & \wp\wp(W)
\end{array}
$$

and that $\hat{V}(p) = \pi^{-1}(V(p))$ for all $p \in \mathcal{P}$ (i.e. π is a strict homomorphism w.r.t. the valuations). The covering is *finite* if all its fibres $\pi^{-1}(w) \subseteq \hat{W}$ are finite.

The definition implies that π induces a global bisimulation between $\mathbb{M}$ and $\hat{\mathbb{M}}$ as $\hat{\mathbb{M}}, \hat{w} \sim \mathbb{M}, \pi(\hat{w})$ for all $\hat{w} \in \hat{W}$.

Towards the desired upgrading arguments, a partial unfolding of the global pattern of equivalence classes $[w]_a$ in an inquisitive $S5$ model can be achieved in finite bisimilar coverings. The eventual covering construction at the global level requires a pre-processing of the local component structures $\mathfrak{M} \restriction [w]_a$ of $\mathsf{I}(\mathfrak{M})$ as introduced in Section 3.2 above:

$$\mathfrak{M} \restriction [w]_a := \big([w]_a, \wp([w]_a), E_a \restriction [w]_a, (P_i \restriction [w]_a)\big),$$

which for a locally full $\mathfrak{M}$ is the full relational encoding of the mono-modal, single-component epistemic model

$$\mathbb{M} \restriction [w]_a = ([w]_a, \Sigma_a \restriction [w]_a, V \restriction [w]_a),$$

whose epistemic inquisitive assignment $\Sigma_a \colon [w]_a \to \wp\wp([w]_a)$ is constant and can be identified with its value $\Sigma_a(w) \in \wp\wp([w]_a)$.

We first explore the extent to which this Σ_a-assignment can be locally modified and simplified in companion models that retain bisimilarity in the global context. This local pre-processing in Section 3.3.1 is specific for epistemic INQML and new here; the techniques for the global pre-processing via local unfoldings in globally bisimilar finite covers, in Section 3.3.2, adapts ideas from modal logics in [26, 14]; both aspects go substantially beyond the treatment of basic INQML in [10].

3.3.1   Local pre-processing.

Since Σ_a is constant across the single equivalence class $[w]_a$, bisimulation – for this single-agent structure – equivalence collapses to its second level $\sim^1$. W.r.t.

28

a fixed set $\mathcal{P}$ of basic propositions each bisimulation type of such a structure is characterised by the collection of propositional types realised overall in its worlds and in the information states in its inquisitive assignment. In particular, multiplicities of these propositional types do not matter at all, neither overall nor in any individual information state: any positive number of realisations is as good as just a single one, so that multiplicities can be uniformly boosted to avoid distinguishing features at the level of FO.

The natural restrictions-cum-reducts $\mathbb{M}\!\restriction\![w]_a$ ignore the inquisitive structure induced by agents $b \neq a$, and fall into one of these global $\sim^1$-equivalence classes. All the remaining complexity of multi-agent inquisitive $S5$ structures arises from the overlapping of a-classes for different $a \in \mathcal{A}$, i.e. from the manner in which the various local structures $\mathbb{M}\!\restriction\![u]_a$ or their relational encodings $\mathfrak{M}\!\restriction\![w]_a$ overlap. In order to freeze the manner in which the local structures on individual a-classes are stitched together, we endow the $\mathbb{M}\!\restriction\![w]_a$ with colourings of their worlds that prescribe their bisimulation types in the global structure $\mathbb{M}$. For different purposes we use different granularities for this colouring. The finest one of interest is the colouring based on the full bisimulation-type in $\mathbb{M}$, with colour set $C = W/\!\sim$, the set of $\sim$-equivalence classes of worlds in $\mathbb{M}$. This is the colouring that needs to be respected in passage to fully bisimilar partner structures like $\hat{\mathfrak{M}}, \hat{w} \sim \mathfrak{M}, w$ in the vertical axes in Figure 4.

Coarser colourings are induced by the $\sim^m$-type, for a fixed finite level m, with colour set $C = W/\!\sim^m$. These may also be obtained via the natural projection that identifies all $\sim$-types that fall into the same $\sim^m$-class. Successive coarsenings of this kind will inform the strategy analysis towards establishing levels of FO-indistinguishability like $\hat{\mathfrak{M}}, \hat{w} \oplus \mathfrak{X} \equiv_r^{(\ell)} \hat{\mathfrak{M}}', \hat{w}' \oplus \mathfrak{X}$ on the basis of $\mathfrak{M}, w \sim^n \mathfrak{M}', w'$ across the horizontal axes in Figure 4. Moreover, the relevant granularity m will be dependent on the distance from the distinguished worlds or states, just as control over levels of bisimilarity decreases with the depth of the exploration. But this granularity m will be uniformly bounded by the initial level n of n-bisimilarity in $\mathfrak{M}, w \sim^n \mathfrak{M}', w'$ or $\mathfrak{M}, s \sim^n \mathfrak{M}', s'$.

Formally, a colouring of worlds by any colour set C can be encoded as a propositional assignment, for new atomic propositions $c \in C$. Unlike propositional assignments in general, these propositions will be mutually exclusive, so as to partition the set of worlds. To reduce conceptual overhead, we encode colourings as functions

$$
\begin{array}{ccc}
\rho\colon W & \to & C \\
w & \mapsto & \rho(w)
\end{array}
\qquad \text{instead of} \qquad
\begin{array}{ccc}
V\colon C & \to & \wp(W) \\
c & \mapsto & V(c) = \{w\colon \rho(w) = c\}
\end{array}
$$

and correspondingly write, e.g., $(W, (\Sigma_a)_{a \in \mathcal{A}}, \rho)$ for the inquisitive epistemic model with the colouring $\rho\colon W \to C$ that assigns to each world its $\sim$-class:

$$
\begin{array}{ccl}
\rho\colon W & \longrightarrow & C := W/\!\sim \\
w & \longmapsto & [w]_\sim = \{u \in W\colon \mathbb{M}, u \sim \mathbb{M}, w\}
\end{array}
$$

The coarser colouring based on $\sim^m$-types is correspondingly formalised in the inquisitive epistemic model $(\mathbb{M}, (\Sigma_a)_{a \in \mathcal{A}}, \rho_m)$, with the colouring $\rho_m\colon W \to$

$W/\sim^m$. Clearly the $\sim$-colouring ρ refines the $\sim^m$-colouring ρ_m, and ρ_n refines ρ_m for $m < n$. In particular, all levels refine ρ_0, which determines each world's original propositional type as induced by V in $\mathbb{M}$.

Finiteness of the set of colours $W/\sim$ for the $\sim$-colouring ρ is guaranteed over each individual finite model $\mathbb{M}$ (or across pairs of finite models $\mathbb{M}$ and $\mathbb{M}'$ to be compared). But also over possibly infinite models the sets of colurs $W/\sim^m$ for the $\sim^m$-colourings ρ_m are unioirmly finitely bounded across all models in a finite signature, for each finite level m.

Definition 3.10. For an a-class $[w]_a$ in an inquisitive model $\mathbb{M}$, let the associated *local $\sim$-structure* be the mono-modal inquisitive $S5$ structure $\mathbb{M}{\restriction}[w]_a, \rho$ with the propositional assignment induced by ρ in $\mathbb{M}$, which colours each world in $[w]_a$ according to its $\sim$-type in $\mathbb{M}$.

For $m \in \mathbb{N}$, let $\mathbb{M}{\restriction}[w]_a, \rho_m$ be the coarsening of granularity m of that local structure: the *local $\sim^m$-structure* with propositional assignment induced by ρ_m, which colours each world in $[w]_a$ by its $\sim^m$-type in $\mathbb{M}$ according to the natural projection $\pi\colon W/\sim \to W/\sim^m$.

Interestingly, a local $\sim$- or $\sim^m$-structure can be modified considerably without changing its bisimulation type. We now want to achieve local structures

- that realise every bisimulation type of worlds locally with high multiplicity; this leads to the notion of *richness* in Definition 3.11;

- whose inquisitive assignment Σ_a is normalised w.r.t. to the structure of the inquisitive state $\Pi = \Sigma_a(w)$ within $\wp([w]_a)$ so as to suppress FO-definable distinctions that are not governed by any level of $\sim^m$; this leads to the notion of *regularity* in Definition 3.13.

Richness. The idea for *richness* is a quantitative one, viz., that all bisimulation types that are realised in any information state in $\Sigma_a(w)$ will be realised with high multiplicity in some superset that is also in $\Sigma_a(w)$. Any desired finite level of richness can be achieved in finite bisimilar coverings (Lemma 3.12), which essentially just put a fixed number of copies of every world.

Definition 3.11. A local structure $\mathbb{M}{\restriction}[w]_a$ in $\mathbb{M}$ is *K-rich* for some $K \in \mathbb{N}$ if, for every information state $s \in \Sigma_a(w)$ there is some s', $s \subseteq s' \in \Sigma_a(w)$, such that every $\sim$-type in $W/\sim$ that is realised in s is realised at least K times in s'. An inquisitive $S5$ structure $\mathbb{M} = (W, (\Sigma_a), V)$ is *K-rich* if each one of its local structures is K-rich, for all $a \in \mathcal{A}$.

Lemma 3.12. *For any $K \in \mathbb{N}$, any inquisitive $S5$ structure $\mathbb{M} = (W, (\Sigma_a), V)$ admits a finite bisimilar covering that is K-rich. If finiteness does not matter, richness can similarly be achieved for any target cardinality.*

Proof. It suffices to take the natural product $\mathbb{M} \times [K]$ of $\mathbb{M}$ with the standard K-element set $[K] = \{1, \ldots, K\}$. This results in a model whose set of worlds is $W \times [K]$, whose propositional assignment is induced by the projection $\pi\colon W \times [K] \to$

W, and whose inquisitive assignment is $\Sigma_a(w, m) := \{s \subseteq W \times [K] : \pi(s) \in \Sigma_a(w)\}$. It is easily checked that $\mathbb{M} \times [K]$ is K-rich and that π is a bisimilar covering in the sense of Definition 3.9. $\qquad \Box$

Regularity. Unlike the merely quantitative flavour of richness, regularity imposes qualitative constraints on the structure of the family of information states in $\Sigma_a(w)$. It requires a modification of the local inquisitive assignment and is not obtained in a covering. The idea is to normalise the downward closed family of information states in $\Pi := \Sigma_a(w)$ within the boolean algebra $\wp([w]_a)$ of subsets of $[w]_a$ as much as possible. This normalisation needs to preserve the bisimulation type as encoded in the colouring ρ. It must also be geared towards passage to the coarser picture induced by increasingly lower levels of m-bisimulation as encoded in the ρ_m.[13] Passage to some suitable degree of regularity should be thought of as a local modification of the representation, rather than the epistemic content, of the inquisitive assignment. Indeed, bisimilarity is rather robust under changes in the actual composition of the inquisitive assignment $\Pi = \Sigma_a(w) \subseteq \wp([w]_a)$ in $\mathbb{M}$. By the inquisitive epistemic nature of $\mathbb{M}$, Σ_a has constant value Π on $[w]_a$ and Π is a downward closed collection of subsets of $[w]_a$ such that $w \in \bigcup \Pi = [w]_a = \sigma_a(w)$. The only additional invariant imposed on Π up to $\sim$ stems from the associated colouring ρ of worlds by their $\sim$-types in $C = W/\sim$; for the inquisitive assignment Π this invariant concerns the colour combinations in

$$\rho(\Pi) := \{\rho(s) : s \in \Pi\} \subseteq \wp(C)$$
$$\text{for } \rho(s) := \{\rho(v) : v \in s\} = \{c \in W/\sim : c \cap s \neq \emptyset\}.$$

Note that $\rho(\Pi)$ is downward closed in $\wp(C)$. Not only the collection of bisimulation types in the set of worlds $[w]_a$, but also in the information states of the inquisitive assignment Π is fully determined by $\rho(\Pi)$. Up to bisimulation, Π can be replaced by any inquisitive epistemic assignment $\hat{\Pi}$ with $\rho(\hat{\Pi}) = \rho(\Pi)$.

Similar remarks apply to finite levels $\sim^m$ relative to the colouring ρ_m with values in $W/\sim^m$, and for analogously defined $\rho_m(\Pi) = \{\rho_m(s) : s \in \Pi\}$. Clearly $\rho_m(\Pi)$ is fully determined by $\rho(\Pi)$, so that $\rho_m(\hat{\Pi}) = \rho_m(\Pi)$ for all admissible variations $\hat{\Pi}$. But crucially, due to the coarsening, colour combinations $\rho_m(s_i)$ for infomation states $s_1, s_2 \in \Pi$ may coincide even if $\rho(s_1) \neq \rho(s_2)$. So the constraint that $\rho(\hat{\Pi}) = \rho(\Pi)$ implies that some colour combinations in $\rho_m(\hat{\Pi}) = \rho_m(\Pi)$ may have to appear in more than one $\subseteq$-maximal element of $\rho_m(\Pi)$ and hence must be realised in at least that many $\subseteq$-maximal elements of any admissible $\hat{\Pi}$.[13]

The assignment $\hat{\Pi}$ we choose will therefore be obtained from an upper bound κ for the possibly unavoidable multiplicity of $\subseteq$-maximal elements in $\rho_m(\Pi)$. So κ is to be adjusted accross the two models to be matched in the upgrading argument of Figure 4; it will always be finite in the context of finite inquisitive models for the finite model theoretic reading of our results, but could necessarily

---

[13]An oversimplification in earlier versions of this preprint wrongly assumed robustness w.r.t. such coarsenings; we are grateful to an anonymous referee for spotting the mistake.
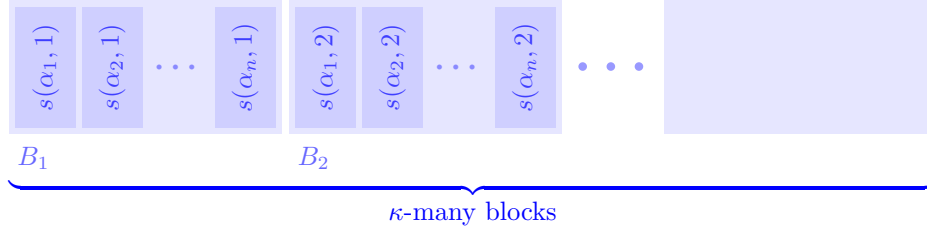
Figure 5: Subdivison of $[w]_a = \sigma_a(w)$ in the local a-structure $\mathbb{M} \restriction [w]_a$ for an inquisitive assignment $\Pi = \Sigma_a(w)$ that is κ-regular at granularity m, where $\rho_m(\Pi \setminus \emptyset) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $\kappa = \{1, 2, \ldots\}$ and each one of the contributions $s(\alpha_j, i)$ in block B_i is one $\subseteq$-maximal element in Π, so that there are precisley κ many such for each available $\sim^m$-type of nonempty information states in Π.

be infinite in the general case of infinite models that might even locally realise infinitely many distinct bisimulation types. Its granularity levels m will be dependent on the distance of the local structures from the distinguished worlds or states in the upgrading scenario of Figure 4. It will always be uniformly bounded by the initial level of n-bisimilarity $\sim^n$ between the two pointed models under consideration. So there is a fixed finite bound on the number of $\sim^m$-colours that may occur. Such an upper bound can uniformly be chosen as a bound on the number of distinct $\sim$-types realised by information states in the inquisitive assignments under consideration.

Towards normalised patterns for $\hat{\Pi} = \hat{\Sigma}_a(w) \subseteq \wp([w]_a)$ we think of decomposing $[w]_a$ into disjoint subsets $B_i \subseteq [w]_a$, which we call *blocks*, such that every nonempty information state $s \in \hat{\Pi} \setminus \{\emptyset\}$ is contained in one of these, and such that each block B_i is a disjoint union of $\subseteq$-maximal information states in $\hat{\Pi}$, one for each $\rho_m(s)$, $s \in \hat{\Pi} \setminus \{\emptyset\}$. In particular $\hat{\Pi}$ will be the union of its restrictions $\hat{\Pi} \restriction B_i := \{s \in \hat{\Pi} \colon s \subseteq B_i\}$ to the κ many blocks, as illustrated in Figure 5.

Definition 3.13. The inquisitive assignment $\Pi = \Sigma_a(w) \subseteq \wp([w]_a)$ in the local structure $\mathbb{M} \restriction [w]_a$ is *κ-regular at granularity m* if $[w]_a = \dot{\bigcup}_{i<\kappa} B_i$ is a disjoint union of precisely κ many blocks $B_i \subseteq [w]_a$ such that $\Pi = \bigcup_i \Pi \restriction B_i$ and

- each of the restrictions $\Pi \restriction B_i$ realises every ρ_m-colour combination $\alpha = \rho_m(s)$ for $s \in \Pi \setminus \{\emptyset\}$ in precisely one $\subseteq$-maximal element $s(\alpha, i) \in \Pi \restriction B_i$;
- these maximal elements $s(\alpha, i) \in \Pi \restriction B_i$ are disjoint for distinct α and form a partition of B_i.

It then follows that Π itself is generated by the combination of the downward closures of the family of disjoint maximal elements $s(\alpha, i)$ across all α and i, which partition $[w]_a = \dot{\bigcup}_{\alpha, i} s(\alpha, i)$.

Lemma 3.14. *Let $\mathbb{M} = (W, (\Sigma_a), V)$ be an inquisitive epistemic model that is sufficiently rich in relation to given parameters κ, K and the cardinality of $W/\sim$; let $m \colon W \to \mathbb{N}$ be any uniformly bounded function prescribing finite levels of granularity for the $\sim^m$-local structures $\mathbb{M} \restriction [w]_a$ at $w \in W$. Then there is a*

globally bisimilar companion $\hat{\mathbb{M}} = (W, (\hat{\Sigma}_a), V)$ *such that* $\hat{\mathbb{M}}, w \sim \mathbb{M}, w$ *for all* $w \in W$ *(so that the* $\sim$*-colouring* ρ *on* W *is unaffected) that is* K*-rich and in which every local structure* $\hat{\mathbb{M}} \restriction [w]_a$ *is* κ*-regular at granularity* $m(w)$.

Proof. It suffices to restructure every $\Pi = \Sigma_a(w) \subseteq \wp([w]_a)$ individually into a $\hat{\Pi}$ that is κ-regular w.r.t. ρ_m for $m = m(w)$ and remains K-rich. Provided that $[w]_a$ has sufficiently many realisations of every $\sim^m$-type (as guaranteed by a sufficiently high level of richness in $\mathbb{M}$), $[w]_a$ can be split into κ many disjoint blocks each of which is split into disjoint $\subseteq$-maximal new information states, precisely one for every $\sim^m$-type $\alpha = \rho(s)$ for $s \in \Sigma_a(w)$, realising every $\sim^m$-type of its worlds at least K times. $\qquad\square$

Observation 3.15. *Every inquisitive epistemic model* $\mathbb{M}$ *admits, for* $K \in \mathbb{N}$, *finite bisimilar coverings* $\hat{\mathbb{M}}$ *that are* K*-rich. As a consequence there are, for given* $K, M \in \mathbb{N}$, *globally bisimilar inquisitive epistemic models* $\hat{\mathbb{M}}$ *sharing the same induced Kripke model* $\mathfrak{K}(\hat{\mathbb{M}})$ *with variations of just the inquisitive assignments, for every given* M*-bounded granularity assignment* $m \colon \hat{W} \to \mathbb{N}$, *that make them both* K*-rich and* κ*-regular w.r.t. that granularity assignment* m. *Such* $\hat{\mathbb{M}}$ *can be chosen to be finite for finite* $\mathbb{M}$ *and* κ.

3.3.2 Global pre-processing

The desired pre-processing needs to unclutter and smooth out the local overlap patterns among local structures $\mathfrak{M} \restriction [w]_a$ in $\mathsf{I}(\mathfrak{M})$ in such a manner that the ℓ-bisimulation type determines the first-order behaviour of ℓ-neighbourhoods up to quantifier rank r (cf. Proposition 3.8). This requires a local unfolding of the connectivity in the underlying Kripke structures and uses ideas developed in [26] to eliminate incidental cycles and overlaps in finite bisimilar coverings — ideas which are applied to plain $S5$ Kripke structures in [14]. Such overlaps, which up to bisimulation equivalence are incidental (i.e. cannot be prescribed by the bisimulation type), occur for instance whenever an a-class and an a'-class for $a \neq a'$ share more than one world. These worlds would be linked by both an R_a-edge and an $R_{a'}$-edge. A bisimilar companion obtained from a classical tree unfolding (by the appropriate closure operations that turn it into an $S5$ frame) would separate these two edges. Similarly any cycle formed by a finite sequence of overlapping a_i-classes that returns to its starting point would not be stable under appropriate bisimilar unfoldings.

For a given epistemic model $\mathbb{M}$ consider the induced Kripke structure $\mathfrak{K} = \mathfrak{K}(\mathbb{M}) = (W, (R_a)_{a \in \mathcal{A}}, V)$, which is a single-sorted, relational multi-modal $S5$ structure with equivalence relations R_a derived from the inquisitive epistemic frame $\mathbb{M}$, with equivalence classes $[w]_a = \sigma_a(w)$. Generic constructions from [26], based on products with suitable Cayley groups, yield finite bisimilar coverings $\pi \colon \hat{\mathfrak{K}} \to \mathfrak{K}$ by $S5$ structures $\hat{\mathfrak{K}}$ that are N-acyclic in the following sense.

Definition 3.16. For $N \geqslant 2$, an epistemic model $\mathbb{M}$ is called *N-acyclic* if the induced $S5$ Kripke frame $(W, (R_a)_{a \in \mathcal{A}})$ underlying $\mathfrak{K} = \mathfrak{K}(\mathbb{M})$ is N-acyclic in the sense that there is no non-trivial cyclic pattern of overlapping a_i-classes of
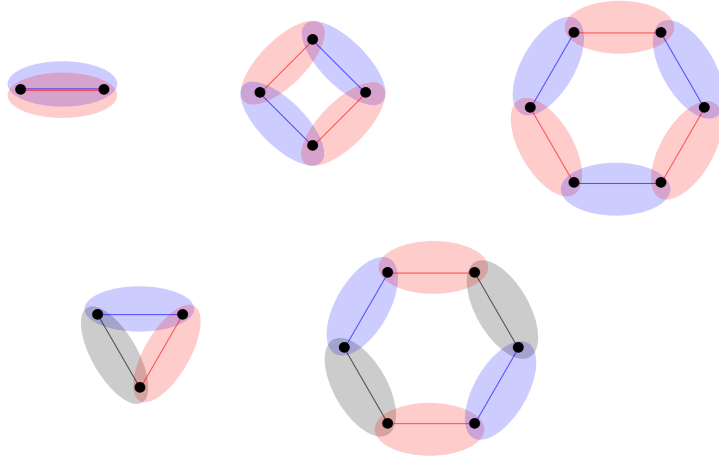
Figure 6: Example of particularly simple unfoldings in bisimilar coverings of $S5$ Kripke frames for agents *red, blue, black*; top row: unfolding of a 2-cycle (overlap of two equivalence classes in two world) into a 4- or 6-cycle; below: unfolding of a 3-cycle into a 6-cycle (just non-trivial equivalence classes colour-coded and worlds in overlaps and links in induced cycles marked).

length up to N: for $2 \leqslant n \leqslant N$, there is no cyclically indexed family $([w_i]_{a_i})_{i \in \mathbb{Z}_n}$ such that $[w_i]_{a_i} = [w_{i+1}]_{a_i}$ while $w_{i+1} \neq w_i$ and $a_{i+1} \neq a_i$.

This acyclicity condition implies in particular that a_i-classes and a_j-classes for agents $a_i \neq a_j$ overlap in at most a single world: if $[w]_{a_1} \cap [w]_{a_2} \neq \{w\}$ the two classes would form a non-trivial 2-cycle. For $N \geqslant 2\ell$ it moreover implies that the overlap pattern between a-local structures is simply tree-like within Gaifman radius ℓ: any two distinct short sequences of such overlaps emanating from the same world cannot meet elsewhere.

By [26], this form of N-acyclicity can be achieved, for any given N, in finite bisimilar coverings in which every a-class is bijectively related to an a-class in the original structure by the projection map of the covering, which induces the bisimulation. We briefly outline the extension of such bisimilar coverings of $\mathfrak{K}(\mathbb{M})$ to bisimilar coverings of the inquisitive epistemic model $\mathbb{M}$. Let $\pi \colon \hat{\mathfrak{K}} \to \mathfrak{K}$ be such an N-acyclic finite bisimilar covering, which in particular means that the graph of the covering projection π is a bisimulation relation. Since a-classes in $\hat{\mathfrak{K}}$ are π-related isomorphic copies of corresponding a-classes in $\mathfrak{K}$, we may consistently endow them with an inquisitive Σ_a-assignment as pulled back from $\mathbb{M}$, to obtain a natural derived finite inquisitive bisimilar covering

$$\pi \colon \hat{\mathbb{M}} \longrightarrow \mathbb{M} = (W, (\Sigma_a), V)$$

in the sense of Definition 3.9. The inquisitive epistemic model $\hat{\mathbb{M}} = (\hat{W}, (\hat{\Sigma}_a), \hat{V})$ is built on the basis of the covering $S5$ Kripke structure $\hat{\mathfrak{K}}$ so that $\mathfrak{K}(\hat{\mathbb{M}}) = \hat{\mathfrak{K}}$. More specifically, for $\hat{w} \in \hat{W}$ consider its $\hat{R}_a$-equivalence class $[\hat{w}]_a \subseteq \hat{W}$, which

is bijectively related by π to the R_a-equivalence class of $w = \pi(\hat{w})$ in $\mathfrak{K}$, $[w]_a \subseteq W$. In restriction to $[\hat{w}]_a$ we put

$$\hat{\Sigma}_a \restriction [\hat{w}]_a \colon [\hat{w}]_a \quad \longrightarrow \quad \wp\wp([\hat{w}]_a)$$
$$\hat{u} \quad \longmapsto \quad \{\pi^{-1}(s) \colon s \in \Sigma_a(w) \subseteq \wp([w]_a)\},$$

where π^{-1} refers to the inverse of the local bijection between $[\hat{w}]_a$ and $[w]_a$ induced by π. In other words, we make the following diagram commute and note that the relevant restriction of π in the right-hand part of the diagram is bijective and part of a bisimulation at the level of the underlying $S5$ Kripke structures, which means in particular that it is compatible with the given propositional assignments:

We merely need to check that the resulting model $\hat{\mathbb{M}}$ is again an inquisitive epistemic model, and that π induces an inquisitive bisimulation, not just a bisimulation at the level of the underlying single-sorted, multi-modal $S5$-structures $\hat{\mathfrak{K}} = \mathfrak{K}(\hat{\mathbb{M}})$ and $\mathfrak{K} = \mathfrak{K}(\mathbb{M})$. That $\hat{\mathbb{M}}$ is an inquisitive epistemic model is straightforward from its construction: $\hat{\Sigma}_a$-values are constant on the a-classes induced by $\hat{R}_a$ just as Σ_a-values are constant across a-classes induced by R_a; and $\hat{w} \in \hat{\sigma}_a(\hat{w})$ follows from the fact that $w \in \sigma_a(w) = [w]_a$.

A non-deterministic winning strategy in the unbounded inquisitive bisimulation on $\hat{\mathbb{M}}, \hat{w}$ and $\mathbb{M}, \pi(\hat{w})$ is given by the following bisimulation relation, to be viewed as a set of winning positions for player **II** in the bisimulation game:

$$Z = \{(\hat{u}, \pi(\hat{u})) \in \hat{W} \times W \colon \hat{u} \in \hat{W}\}$$
$$\cup \{(\hat{s}, \pi(\hat{s})) \in \wp(\hat{W}) \times \wp(W) \colon a \in \mathcal{A}, \hat{w} \in \hat{W}, \hat{s} \in \hat{\Sigma}_a(\hat{w})\},$$

which is the natural lifting of π to information states. It is easily verified that player **II** can stay in Z in response to any challenges by player **I** (Z satisfies the usual back&forth conditions). We summarise these findings as follows.

Lemma 3.17. *Any inquisitive epistemic model* $\mathbb{M} = (W, (\Sigma_a), V)$ *admits, for every* $N \in \mathbb{N}$*, a finite bisimilar covering of the form*

$$\pi \colon \hat{\mathbb{M}} \longrightarrow \mathbb{M}$$

by an inquisitive epistemic model $\hat{\mathbb{M}} = (\hat{W}, (\hat{\Sigma}_a), \hat{V})$ *such that*

(i) $\hat{\mathbb{M}}$ *is N-acyclic in the sense of Definition 3.16;*

(ii) *the global bisimulation induced by π is an isomorphism in restriction to each local structure $\hat{\mathbb{M}} \restriction [\hat{w}]_a$ of $\hat{\mathbb{M}}$, which is isomorphically mapped by π onto the local a-structure $\mathbb{M} \restriction [\pi(\hat{w})]_a$ of $\mathbb{M}$.*

It is noteworthy that condition *(ii)* guarantees that degrees of regularity and richness are preserved in the covering, simply because they are properties of the local structures $\mathbb{M} \restriction [w]_a$, which up to isomorphism are the same in $\mathbb{M}$ and in $\hat{\mathbb{M}}$. So, in combination with Observation 3.15, we obtain the following.

Proposition 3.18. *For any desired values of $K, N \in \mathbb{N}$ and (finite or infinite) cardinality κ, and for any finite bound M on relevant granularities, any world- or state-pointed inquisitive epistemic model $\mathbb{M}$ admits a globally bisimilar N-acyclic companions $\hat{\mathbb{M}} \sim \mathbb{M}$, all based on the same N-acyclic induced Kripke model $\mathfrak{K}(\hat{\mathbb{M}})$, whose epistemic assignments can be chosen K-rich and κ-regular w.r.t. any given uniformly bounded granularity assignment $m \colon \hat{W} \to \mathbb{N}$. Such $\hat{\mathbb{M}}$ can be chosen finite for finite $\mathbb{M}$ and κ. In the non-trivially state-pointed case of $\mathbb{M}, s$, companions $\mathbb{M}_s$ based on the dummy agent s admit bisimilar companions $\hat{\mathbb{M}}, \hat{s} \sim \mathbb{M}_s, s$ with these same properties.*

The non-trivially state-pointed case may deserve further comment. The desired $\hat{\mathbb{M}}$ in this case is based on the augmentation $\mathbb{M}_s$ with the extra agent s. The local pre-processing produces in particular a rich and κ-regular s-local structure over a set of worlds $\hat{s}$; N-acyclicity for $N > 2\ell$ in the global pre-processing then serves to scatter the worlds in $\hat{s}$ within $\mathfrak{K}(\hat{\mathbb{M}})$ w.r.t. $\mathcal{A}$ in the sense that no two $(E_a)_{a \in \mathcal{A}}$-paths of length up to ℓ from distinct worlds in $\hat{s}$ can meet: they are directly E_s-linked and N-cyclic patterns are ruled out.

3.4 Proof of the characterisation theorem

In the remainder of this section we complete the proof of our main theorem, Theorem 1.2, the precise statement of which was given in Section 3.1 on page 16.

The proof of this theorem has, through preparations in the preceding sections, already been reduced to the upgrading target in equation (6). For the world-pointed case:

$$\mathsf{I}(\hat{\mathfrak{M}}, \hat{w}) \;\equiv_r^{(\ell)}\; \mathsf{I}(\hat{\mathfrak{M}}', \hat{w}'),$$

where we now take $\hat{\mathfrak{M}}$ and $\hat{\mathfrak{M}}'$ to be locally full relational encodings of bisimilar companions of any given pair of $\sim^n$-bisimilar pointed models $\mathbb{M}(\mathfrak{M})$ and $\mathbb{M}(\mathfrak{M}')$, which are obtained through the detour summarised in Proposition 3.18, for suitable values of the parameters of K and N for richness and acyclicity, and κ and m for regularity. The parameters of the target equivalence, ℓ and r, are determined by the quantifier rank of the given $\sim$-invariant $\varphi \in \mathsf{FO}$ that is to be shown to be $\sim^n$-invariant, for a value of n still to be determined by the Ehrenfeucht–Fraïssé argument for $\equiv_r^{(\ell)}$.

Recall from our preliminary discussion at the end Section 3.2 how the desired upgrading argument for the crucial compactness claim in line with Proposition 3.8, calls for a strategy in the r-round first-order Ehrenfeucht–Fraïssé game

over the ℓ-neighbourhoods of exploded views of the pre-processed models. Due to the nature of the exploded views as conglomerates of interconnected a-local structures, the desired strategy analysis focuses on a separation between

(a) the global analysis of the FO-game in the ℓ-neighbourhoods, now involving a tree-like connectivity pattern between a-local component structures in sufficiently acyclic models.

(b) the a-local concerns, now involving sufficiently rich and suitably regular models, where the FO-game over these two-sorted structures can be analysed in terms an MSO-analysis over their first sorts, to be treated first, in Section 3.4.1.

The granularity for the regularity requirements in (b) will depend on the distance from the distinguished world or state, since control over finite levels of bisimulation equivalence decreases with this distance from the root in the tree-like core structure. The resulting strategy advice for player **II** is brought together in Section 3.4.2. We first turn to these a-local component structures.

3.4.1 MSO-equivalence over local structures

Recall Definition 3.10 of local structures $\mathbb{M} \restriction [w]_a$ with colouring ρ_m that marks out $\sim^m$-types in the surrounding $\mathbb{M}$ or $\mathbb{M}'$. In the locally full relational encoding $\mathfrak{M} = \mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$, we correspondingly consider $\mathfrak{M} \restriction [w]_a, \rho_m$ as a mono-modal relational model. We apply Ehrenfeucht–Fraïssé techniques for MSO in order to obtain guarantees for levels of first-order equivalence $\equiv_k$ over these two-sorted local $\sim^m$-structures, assuming a suitable degree of richness and regularity w.r.t. ρ_m. Richness matters since FO can assess cardinality distinctions between information states in the second sort, but just up to certain finite thresholds. These thresholds are inherited from MSO-definability over plain coloured sets over the first sort.

To compare the sizes of sets P, P' up to some threshold $d \in \mathbb{N}$, we introduce this cardinality equality with cut-off:

$$|P| =_d |P'| \quad :\Leftrightarrow \quad |P| = |P'| \text{ or } |P|, |P'| \geqslant d.$$

For the analysis of a set W with any C-colouring $\rho \colon W \to C$, we also use its representation as a plain relational structure $(W, (W_c)_{c \in C})$ with colour classes W_c viewed as the interpretations of unary relations (monadic predicates). So we interchangeably think of representations of W, ρ as a set with a colouring $\rho \colon W \to C$ or as a relational structure partitioned into these colour classes. In our applications, W will always be the first sort $[w]_a$ of some local $\sim^m$-structure, coloured by ρ_m with the finite colour set for $\sim^m$-types of worlds.

With any subset $P \subseteq W$ we correspondingly associate the family

$$(P_c)_{c \in C} \quad \text{where} \quad P_c := \{w \in P \colon \rho(w) = c\} = P \cap W_c.$$

For the comparison of $P \subseteq W$ with $P' \subseteq W'$ from two C-coloured sets W, ρ and W', ρ' we define the equivalence relation $\approx_d^C$ according to

$$P \approx_d^C P' \quad \text{iff} \quad |P_c| =_d |P_c'| \text{ for all } c \in C;$$

and generalise this to k-tuples of subsets $\mathbf{P} \in (\wp(W))^k$ and $\mathbf{P}' \in (\wp(W'))^k$ by putting

$$\mathbf{P} \approx_d^C \mathbf{P}' \quad \text{iff} \quad \zeta(\mathbf{P}) \approx_d^C \zeta(\mathbf{P}') \text{ for all boolean terms } \zeta(\mathbf{X}).^{14}$$

The following is folklore, but we indicate the straightforward proof below.

Lemma 3.19. *For C-coloured sets W, W' with k-tuples $\mathbf{P} \in \wp(W)^k$ and $\mathbf{P}' \in \wp(W')^k$ of subsets, and for $d = 2^q$:*

$$\mathbf{P} \approx_d^C \mathbf{P}' \quad \Rightarrow \quad (W, (W_c)_{c \in C}, \mathbf{P}) \equiv_q^{\mathsf{MSO}} (W', (W'_c)_{c \in C}, \mathbf{P}').$$

Note that assignments to first-order element variables x are implicit in the form of singleton sets for $\{x\}$ (which induces but a small constant shift in quantifier rank levels).

Proofsketch. The proof is by induction on q, and we first look at the case with trivial colouring (all elements are of the same colour). For the induction step assume that $\mathbf{P} \approx_{2d}^C \mathbf{P}'$ and suppose w.l.o.g. that the first player proposes a subset $P \subseteq W$ so that the second player needs to find a response $P' \subseteq W'$ such that $\mathbf{P}P \approx_d^C \mathbf{P}'P'$. Decompose P and its complement $\bar{P}$ into their intersections with the atoms of the boolean algebra generated by $\mathbf{P}$ in $\wp(W)$. Then each part of these partitions of P and $\bar{P}$ can be matched with a subset of the corresponding atoms of the boolean algebra generated by $\mathbf{P}'$ in $\wp(W')$ in such a manner that corresponding parts match in the sense of $=_d$. This just uses the assumption that $|\zeta(\mathbf{P})| =_{2d} |\zeta(\mathbf{P}')|$ for each ζ. In the case of non-trivial C-colourings, the same argument can be played out within each pair of colour classes W_c and W'_c after decomposition of P and $\bar{P}$ into their C-partitions, and recombination of P' as the union of the matching P'_c obtained for the P_c. □

We next transfer this MSO-equivalence between ρ_m-coloured first sorts of suitable pre-processed local component structures to FO-equivalence between the two-sorted relational encodings of these local $\sim^m$-structures $\mathfrak{M} \restriction [w]_a, \rho_m$ and $\mathfrak{M}' \restriction [w]_a, \rho_m$ that locally realise the same $\sim^m$-types of information states and are κ-regular at this granularity m for the same value of κ.

Towards an interpretation of the 2-sorted full relational encodings of local $\sim^m$-structures like $\mathfrak{M} \restriction [w]_a, \rho_m$ in terms of MSO-definability over $[w]_a$, it is necessary to make available the information states in $\Pi := \Sigma_a(w)$. Due to κ-regularity at granularity m, this reduces to marking out the $\subseteq$-maximal informatiobn states $s(\alpha, i)$ in each block B_i, for every one of the finitely many colour combinations $\alpha = \rho_m(s) \in \rho_m(\Pi) \setminus \{\emptyset\}$ for non-trivial information states s realised in $\mathfrak{M} \restriction [w]_a, \rho_m$. To this end we look at MSO-interpretability over derived single-sorted structures over $[w]_a$. These encode the block decomposition (by

---

[14]Here a *boolean term* is any term in the functional language of Boolean algebras, with binary operations for union and intersection and a unary operation for complementation.

an equivalence relation B) and the $\subseteq$-maximal constituents across the blocks (by finitely many unary predicates P_α):

$$\mathfrak{b}_m([w]_a) := ([w]_a, \rho_m, B, (P_\alpha))$$
$$\text{for } B := \bigcup_{i<\kappa} B_i^2 = \{(u,v) \in [w]_a^2 : u, v \text{ in the same block}\},$$
$$P_\alpha := \bigcup_{i<\kappa} s(\alpha, i).$$

The full relational encoding of a $\sim^m$-local structure $\mathfrak{M} \restriction [w]_a, \rho_m$ becomes MSO-interpretable over these $\mathfrak{b}_m([w]_a)$. In particular, for $s \in \wp([w]_a)$ we have $s \in \Sigma_a(w)$ if, and only if

$$\mathfrak{b}_m([w]_a), s \models \exists X \Big(s \subseteq X \wedge \forall x \forall y \big((Xx \wedge Xy) \to Bxy\big) \wedge \bigvee_\alpha X \subseteq P_\alpha \Big).$$

It follows that, for some fixed set-off c stemming from the underlying interpretation, and tuples $\mathbf{s}, \mathbf{s}'$ of set parameters over $[w]_a$ and $[w']_a$,

$$\mathfrak{b}_m([w]_a), \mathbf{s} \equiv_{r+c}^{\mathsf{MSO}} \mathfrak{b}_m([w']_a), \mathbf{s}' \;\Rightarrow\; \mathfrak{M} \restriction [w]_a, \rho_m, \mathbf{s} \equiv_r^{\mathsf{FO}} \mathfrak{M}' \restriction [w']_a, \rho_m, \mathbf{s}' \quad (7)$$

In the analysis of FO-equivalences between local $\sim^m$-structures (as part of the Ehrenfeucht–Fraïssé argument over exploded views as in the upgrading as in Figure 4) we shall encounter choices of elements inside these 2-sorted structures as parameters. These may be worlds $v \in [w]_a$ (first sort) or subsets $s \subseteq [w]_a$ (second sort). Looking at these in the MSO-perspective of the above interpretation, we can treat both kinds as monadic second-order parameters by associating $v \in [w]_a$ with $\{v\} \subseteq [w]_a$. In particular, a source world $w \in [w]_a$ for the local $\sim^m$-structure associated with $[w]_a$ can be fed as a distinguished world in $\mathfrak{M} \restriction [w]_a, \rho_m$ and as a singleton set parameter $\{w\} \subseteq [w]_a$ in the MSO-analysis of $\mathfrak{b}_m([w]_a)$.

Lemma 3.20. *Any pair of m-bisimilar pointed $\sim^m$-local structures $\mathbb{M} \restriction [w]_a, w$ and $\mathbb{M}' \restriction [w']_a, w'$ for $m \geqslant 1$ that are both κ-regular for the same (finite or infinite) value of κ, realise the same colour combinations $\alpha \subseteq \rho_m([w]_a) = \rho_m([w']_a$ as nonempty contributions to the partitions of $[w]_a = \dot{\bigcup}_\alpha P_\alpha$ and $[w']_a = \dot{\bigcup}_\alpha P'_\alpha$, respectively. Moreover*

$$\mathfrak{b}_m([w]_a), \{w\} \equiv_q^{\mathsf{MSO}} \mathfrak{b}_m([w']_a), \{w'\},$$

provided both are sufficiently rich in relation to q.

It then follows from equation 7 that in this situation also

$$\mathfrak{M} \restriction [w]_a, \rho_m, w \equiv_k^{\mathsf{FO}} \mathfrak{M}' \restriction [w']_a, \rho_m, w',$$

given the appropriate level of richness for $q = k + c$ (cf. Corollary 3.21 below).

Proof. Note that in the degenerate case of $\kappa = 1$ (regularity in a single block, trivial B), the claim follows directly from Lemma 3.19. The extension to finite κ then follows from compatibility of $\equiv^{\mathsf{MSO}}$ with disjoint unions of structures

(in application to matched blocks B_i and B_i' as marked out by the equivalence relations B and B'). But also in the general case of possibly infinite κ, any choice of a bijection between the κ many blocks of either side, starting with the blocks B_0 and B_0' that contain the distinguished worlds w and w', respectively, allows us to put together a response to a challenge by the first player from responses in the restrictions of the games over matched pairs of blocks, just as in the proof sketch for Lemma 3.19, but w.r.t. a decomposition into a possibly infinite number of non-trivial restrictions to individual blocks. $\qquad\square$

In the comparisons between κ-regular local $\sim^m$-structures $\mathbb{M}\!\restriction\![w]_a, \rho_m$ and $\mathbb{M}'\!\restriction\![w']_a, \rho_m$ it is essential that $m \geqslant 1$. As remarked in the lemma, $\sim^1$ between any two distinguished worlds, $\mathbb{M}\!\restriction\![w]_a, \rho_m, w \sim^1 \mathbb{M}'\!\restriction\![w']_a, \rho_m, w'$, implies that these two local structures are globally bisimilar and in particular realise the same $\sim^m$-types both of worlds $v \in [w]_a$ and $v' \in [w']_a$ and of information states $s \in \Pi := \Sigma_a(w)$ and $s \in \Pi' := \Sigma_a'(w)$. The encodings of the block structure and $\subseteq$-maximal information states in the κ-regular inquisitive assignments Π and Π' were essential for the background MSO-analysis over the first sorts, but we use these results towards the FO-analysis over their 2-sorted full relational encodings. In these terms we have established the following.

Corollary 3.21. *If $\mathbb{M}$ and $\mathbb{M}'$ are both κ-regular (for the same κ) at granularity m and sufficiently rich (in relation to $k \in \mathbb{N}$), then the following holds for any two worlds $v \in [w]_a$ and $v' \in [w']_a$:*

$$\mathbb{M}\!\restriction\![w]_a, \rho_m, v \sim^1 \mathbb{M}'\!\restriction\![w']_a, \rho_m, v' \;\Rightarrow$$
$$\mathbb{M}\!\restriction\![w]_a, \rho_m, v \sim \mathbb{M}'\!\restriction\![w']_a, \rho_m, v' \;\Rightarrow\; \mathfrak{M}\!\restriction\![w]_a, \rho_m, v \equiv_k \mathfrak{M}\!\restriction\![w']_a, \rho_m, v'.$$

So Observation 3.15 and Proposition 3.18, any two inquisitive epistemic models $\mathbb{M}$ and $\mathbb{M}'$ admit, for any target value $N, k \in \mathbb{N}$, globally bisimilar N-acyclic companions $\hat{\mathbb{M}} \sim \mathbb{M}$ and $\hat{\mathbb{M}}' \sim \mathbb{M}'$, finite if $\mathbb{M}$ and $\mathbb{M}'$ are finite, whose locally full relational encodings satisfy the above transfer within every one of their local $\sim^m$-structures. This conditional guarantee of FO-indistinguishability is good for any allocation of the granularities $m \geqslant 1$, which moreover may take different values for different pairs of local structures that are to be matched. It is important that the local analysis reflected in the corollary obviously imports – through ρ_m – global information about $\sim^m$-types in $\hat{\mathbb{M}}$ and $\hat{\mathbb{M}}'$ into the local structures at hand, but that it does not rely on a uniform allocation of granularity across $\hat{\mathbb{M}}$ or $\hat{\mathbb{M}}'$.

3.4.2 Local FO-equivalence of exploded views

Returning to the global view of the underlying models $\mathbb{M} = (W, (\Sigma_a)_{a \in \mathcal{A}}, V)$ and $\mathbb{M}' = (W', (\Sigma_a')_{a \in \mathcal{A}}, V')$ and their exploded views, and assuming $\mathbb{M}, w \sim^n \mathbb{M}', w'$, we recall equation (6) and its rôle on the lefthand side of Figure 4 for the world-pointed case. We first cover this basic, world-pointed version. The adaptation to the non-trivially state-pointed variant, involving the dummy agent s, will then be straightforward. For the world-pointed version, we need to lift the

$[\,\cdot\,]_a$-local equivalences (at suitable granularities $1 \leqslant m \leqslant n$) in matching local $\sim^m$-structures to the ℓ-local equivalence

$$\mathsf{I}(\hat{\mathfrak{M}}, \hat{w}) \equiv^{(\ell)}_r \mathsf{I}(\hat{\mathfrak{M}}', \hat{w}')$$

around the distinguished worlds $\hat{w}$ and $\hat{w}'$. We aim to employ the criterion of N-acyclicity from Definition 3.16, more specifically its consequences for connectivity patterns in the exploded views. In order to reduce the notational overhead in the following we assume w.l.o.g. that $\hat{\mathfrak{M}} = \mathfrak{M}$ and $\hat{\mathfrak{M}}' = \mathfrak{M}'$, or that $\mathbb{M}$ and $\mathbb{M}'$ themselves have been pre-processed in the sense of Proposition 3.18. In the exploded view of such N-acyclic $\mathbb{M}$, any two local structures $\mathfrak{M} \restriction [w_1]_{a_1} \neq \mathfrak{M} \restriction [w_2]_{a_2}$ can be linked through at most one element v shared by $[w_1]_{a_1}$ and $[w_2]_{a_2}$. This follows from Definition 3.16 for any $N \geqslant 2$. Such a link, if it exists, relates the copies of v in the two component structures of $\mathsf{I}(\mathfrak{M})$ by a path of length 2, which is formed by two R_I-edges emanating from the node $v \in W$ in the core part. So this core part $(W, (E_a)_{a \in \mathcal{A}})$, which is the underlying $S5$ frame, provides the only links between individual local constituents $\mathfrak{M} \restriction [w_i]_{a_i}$ in the exploded view $\mathsf{I}(\mathbb{M})$. For N-acyclic $\mathbb{M}$, this core part is N-acyclic, so that there cannot be any non-trivial overlaps between any two distinct threads of lengths up to ℓ of local constituents $\mathfrak{M} \restriction [w_i]_{a_i}$ in the exploded view $\mathsf{I}(\mathbb{M})$ emanating from the same $v \in W$, provided $N > 2\ell$.

Lemma 3.22. *Let $\mathbb{M}$ and $\mathbb{M}'$ be N-acyclic for $N > 2\ell$, κ-regular (for the same κ) at granularity m (depending on depth in the tree-like exploded views, i.e. on distance from the distinguished worlds w, w') and sufficiently rich to support the claim of Corollary 3.21. Then their exploded views $\mathsf{I}(\mathbb{M}) = \mathsf{I}(\mathfrak{M}^{\mathsf{lf}}(\mathbb{M}))$ and $\mathsf{I}(\mathbb{M}') = \mathsf{I}(\mathfrak{M}^{\mathsf{lf}}(\mathbb{M}'))$ satisfy the following for any pair of worlds such that $\mathbb{M}, w \sim^n \mathbb{M}', w'$ for $n \geqslant \ell + 1$:*

$$\mathsf{I}(\mathbb{M}), w \equiv^{(\ell)}_r \mathsf{I}(\mathbb{M}'), w' \quad i.e. \quad \mathsf{I}(\mathbb{M}) \restriction N^\ell(w), w \equiv_r \mathsf{I}(\mathbb{M}') \restriction N^\ell(w'), w'.$$

The proof of this technical core lemma is given in terms of a winning strategy for player **II** in the r-round game on the pointed ℓ-neighbourhoods of w and w' in $\mathsf{I}(\mathbb{M})$ and $\mathsf{I}(\mathbb{M}')$. As outlined at the start of this section, Section 3.4, such a strategy can be based on a combination of strategies in

(a) the r-round first-order game on the tree-like underlying $S5$ Kripke structures on the first sort (worlds) with a-classes for accessibility and a colouring of worlds that reflects sufficient finite approximations to their bisimulation types in $\mathbb{M}$, which in particular suffice to guarantee 1-bisimulation equivalence and hence also $\equiv_r$ of local component structures of appropriate granularities;[15]

(b) local strategies in the r-round games on the relational encodings of the local component structures (with colourings at the appropriate level of granularity) induced by pebbled pairs of worlds or states. These strategies are ultimately based on local MSO-games over their first sorts, as discussed in connection with Corollary 3.21.

---

[15]It is important here that just certain finite levels of bisimilarity, not the full $\sim$-type, can be controlled.

The requirements w.r.t. $\sim^m$-matches will decrease, in terms of the parameter m, with distance from the source worlds w and w', but keeping $m \geqslant 1$ to uniformly support (a).

Proof of Lemma 3.22. Let $\mathfrak{N} := \mathsf{I}(\mathbb{M}) \restriction N^\ell(w)$ and $\mathfrak{N}' := \mathsf{I}(\mathbb{M}') \restriction N^\ell(w')$; for the local component structures we directly refer to the locally full relational encodings $\mathfrak{M} := \mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$ and $\mathfrak{M}' := \mathfrak{M}^{\mathsf{lf}}(\mathbb{M}')$. Regularity in matching local structures $\mathfrak{M} \restriction [u]_a$ and $\mathfrak{M}' \restriction [u']_a$ is imposed at granularity $m(u) := \ell - d(w, u) + 1 = \ell - d(w', u') = m(u')$ where d is the depth of u, u' in the tree-like core parts of the exploded views, i.e. distance from w, w'.

It follows that $1 \leqslant m(u), m(u') \leqslant \ell + 1 \leqslant n$ is uniformly bounded by the value n in the initial assumption about $\mathbb{M}, w \sim^n \mathbb{M}', w'$.

The strategy analysis in support of the lemma is provided in three steps:

(I) parameterisation of game positions player **II** may have to deal with;

(II) abstraction of an invariant capturing the relevant information in those;

(III) showing how player **II** can maintain this invariant through r rounds.

(I) Parameterisation of game positions. We consider any position $\mathbf{u}; \mathbf{u}'$ in the r-round first-order game on $\mathfrak{N}$ and $\mathfrak{N}'$, starting with pebbles on w and w' which are equivalent w.r.t. $\sim^{\ell+1}$ in $\mathbb{M}$ and $\mathbb{M}'$. It will be of the form $\mathbf{u} = (u_0, \ldots, u_k)$ and $\mathbf{u}' = (u_0', \ldots, u_k')$ for some $k \leqslant r$ with $u_0 = w$, $u_0' = w'$, and, for $1 \leqslant i \leqslant k$, either

(i) $u_i = w_i$ and $u_i' = w_i'$ are elements of the first sort in the core parts that are linked to w and w', respectively, by unique paths of overlapping a-classes, or

(ii) $u_i = (w_i)_a$ and $u_i' = (w_i')_a$ are elements of the first sort in local constituents $[w_i]_a$ and $[w_i']_a$, for some $a = a_i \in \mathcal{A}$ and w_i, w_i' in the core parts as in (i), or

(iii) $u_i = s_i$ and $u_i' = s_i'$ are of the second sort (information states) in some $\wp([w_i]_a)$, respectively $\wp([w_i']_a)$, for some $a = a_i \in \mathcal{A}$ and a pair of uniquely determined constituents $[w_i]_a$ and $[w_i']_a$, for w_i, w_i' otherwise as in (i).

In cases (ii) and (iii) the $a \in \mathcal{A}$ is the same in both structures, or else player **I** would have won. We think of the w_i and w_i' in the core parts as the anchors of actual choices made in corresponding rounds, which may additionally involve an $a_i \in \mathcal{A}$ and subsets s_i and s_i' from the second sorts. We note that in case (iii) these anchor worlds w_i, w_i' are not uniquely determined as representatives of their a_i-classes; but we shall argue below that matching choices of these anchors can be made. So we describe the game position after k rounds, for $k \leqslant r$, in terms of the associated $\mathbf{w} = (w_0, \ldots, w_k)$ and $\mathbf{w}' = (w_0', \ldots, w_k')$ and (for some i) additional a_i (as in (ii)) or (a_i, s_i) and (a_i, s_i') (as in (iii)).

We augment the content of the current position further to an invariant that is to be maintained by player **II**, as follows.

(II) Abstraction of an invariant. With the tuple $\mathbf{w}$ of anchor worlds in the core part of $\mathfrak{N}$ we associate a tree-like hypergraph structure whose hyperedges represent overlapping a-classes in the underlying basic modal $S5$-frame. For this we use a minimal spanning sub-tree containing the worlds listed in $\mathbf{w}$. We write $\mathrm{tree}(\mathbf{w})$ for this tree structure. For its vertices $\mathrm{tree}(\mathbf{w})$ comprises, for every w_i in $\mathbf{w}$, the unique sequence of worlds in which those a-classes intersect that make up the shortest connecting path from the root w to w_i, as well as its end points, $w = w_0$ and w_i. The hyperedges of $\mathrm{tree}(\mathbf{w})$ are formed by the non-trivial subsets of its vertices that fall within the same a-class, labelled by the corresponding a (any such subset forms a clique w.r.t. the corresponding R_a, for a unique $a \in \mathcal{A}$ due to acyclicity). So far the w_i are covered by a unique representation of the overlapping a-classes along these shortest connecting paths to the root. We label each vertex u of this tree structure that stems from a world of $\mathbb{M}$ by its $\sim^{\ell-d+1}$-type in $\mathbb{M}$, where $d = d(w, u)$ is its distance from the root $w = w_0$ (its depth in the tree structure); and label its hyperedges by the appropriate agents $a \in \mathcal{A}$. The tree $\mathrm{tree}(\mathbf{w}')$ on the side of $\mathfrak{N}'$ is similarly defined. After k rounds, each of these trees can have at most $1 + k\ell$ vertices and each individual hyperedge can have at most $k + 1$ vertices. For worlds u of $\mathbb{M}$ at distance $d = d(w, u)$, agent $a \in \mathcal{A}$ and $m \in \mathbb{N}$, we now denote as $\mathfrak{M} \restriction [u]_a, \rho_{\sim^m}$ the locally full (i.e. full) relational encoding of the local $\sim^m$-structure $\mathbb{M} \restriction [u]_a, \rho_{\sim^m}$ at granularity $m = \ell - d + 1$, which has colours for $\sim^m$-types in $\mathbb{M}$ (cf. Definition 3.10); and analogously on the side of $\mathfrak{N}'$. We argue that player **II** can maintain the following isomorphy conditions in terms of these tree structures $\mathrm{tree}(\mathbf{w})$ and $\mathrm{tree}(\mathbf{w}')$ through r rounds, and thus force a win.

$\mathsf{Inv}(k)$ after round k:

(1) the tree structures $\mathrm{tree}(\mathbf{w})$ and $\mathrm{tree}(\mathbf{w}')$ spanned by $\mathbf{w} = (w_0, \ldots, w_k)$ and $\mathbf{w}' = (w'_0, \ldots, w'_k)$ are isomorphic (as hyperedge- and vertex-labelled hypergraphs) via an isomorphism ζ that maps $\mathbf{w}$ to $\mathbf{w}'$:

$$\zeta \colon \mathrm{tree}(\mathbf{w}), \mathbf{w} \simeq \mathrm{tree}(\mathbf{w}'), \mathbf{w}'$$

(2) for each pair of worlds $u \in \mathrm{tree}(\mathbf{w})$ and $u' \in \mathrm{tree}(\mathbf{w}')$ that are linked according to $u' = \zeta(u)$ in (1) at depth $d = d(w, u) = d(w', u')$, and for any label a:

$$\mathfrak{M} \restriction [u]_a, \rho_{\sim^m}, \mathbf{s} \equiv_{r-z+1} \mathfrak{M}' \restriction [u']_a, \rho_{\sim^m}, \mathbf{s}' \quad \text{for } m = \ell - d + 1,$$

where $\mathbf{s}$ and $\mathbf{s}'$ are tuples of size z that coherently list any singleton information states corresponding to the tree vertices incident with that a-hyperedge and any information states $s_i \subseteq [u]_a = \sigma_a(u)$ and $s'_i \subseteq [u']_a = \sigma_a(u')$ that may have been chosen in the second sort of the corresponding local component structures during the first k rounds of the game.

(III) Maintaining the invariant. In the form of $\mathsf{Inv}(0)$ these conditions are satisfied at the start of the game: (2) in this case does not add anything beyond (1), which in turn is a consequence of the assumption that $\mathbb{M}, w_0 \sim^n \mathbb{M}', w'_0$.

We show how to maintain the invariant (conditions (1) and (2)) through round k, in which player **I** may either pebble an element of the first sort, i.e. a world in the core part (w_k or w'_k) or in one of the local component structures ($(w_k)_a$ or $(w'_k)_a$), $a = a_k$, or an element of the second sort, i.e. an information state (s_k in some $[w_k]_a$ or s'_k in some $[w'_k]_a$, $a = a_k$,). We refer to the position before this round by parameters $\mathbf{w}, \mathbf{w}', \mathrm{tree}(\mathbf{w}), \ldots$ as described above, but at level $k-1$. Assuming $\mathsf{Inv}(k-1)$ we need to show how player **II** can find responses so as to satisfy $\mathsf{Inv}(k)$. By symmetry we may assume that **I** opens the k-th round by putting the next pebble on the side of $\mathfrak{N}$.

Case 1. A move in the second sort: **I** chooses an information state s_k so that $s_k \subseteq \wp([w_k]_a)$ for some uniquely determined local component $\mathfrak{M} \restriction [w_k]_a$, $a = a_k$.

Case 1.1: if $[w_k]_a = [u]_a$ for some vertex u in $\mathrm{tree}(\mathbf{w})$, u at depth d say, then we look, for $u' = \zeta(u)$, at one round in the game for

$$\mathfrak{M} \restriction [u]_a, \rho_{\sim^m}, \mathbf{s} \equiv_{r-z+1} \mathfrak{M}' \restriction [u']_a, \rho_{\sim^m}, \mathbf{s}' \quad \text{for } m = \ell - d + 1$$

with a move by the first player on s_k, which extends $\mathbf{s}$ to $\mathbf{s}s_k$; this has an adequate response s'_k for the second player, which extends $\mathbf{s}'$ to $\mathbf{s}s'_k$ and guarantees $\equiv_{r-z}$ (this is the appropriate level since the tuples $\mathbf{s}$ and $\mathbf{s}'$ have been extended by one component).

Case 1.2: $[w_k]_a$ is "new" and the appropriate w'_k that satisfies conditions (1) and (2) has to be located in a first step that simulates a move on $(w_k)_a$ (treated as Case 3), after which we may proceed as in Case 1.1.

Case 2. Suppose **I** chooses a world in the core part of $\mathfrak{N}$ in the k-th round, say w_k. The choice of an appropriate match w'_k is treated by induction on the distance that the newly chosen world w_k has from $\mathrm{tree}(\mathbf{w})$. In the base case, distance 0 from $\mathrm{tree}(\mathbf{w})$, w_k is a vertex of $\mathrm{tree}(\mathbf{w})$ and nothing needs to be updated: the response is dictated by the existing isomorphism ζ according to (1). In all other cases, $\mathrm{tree}(\mathbf{w})$ and ζ need to be extended to encompass the new w_k. The new world w_k can be joined to $\mathrm{tree}(\mathbf{w})$ by a unique shortest path of overlapping a-classes of length greater than 0 that connects it to $\mathrm{tree}(\mathbf{w})$. The new branch in $\mathrm{tree}(\mathbf{w})$ will be joined to $\mathrm{tree}(\mathbf{w})$ either through a new a-hyperedge emanating from an existing vertex u of $\mathrm{tree}(\mathbf{w})$ (treated in Case 2.1) or through a new vertex u that is part of an a-class that already contributes an existing hyperedge of $\mathrm{tree}(\mathbf{w})$ (treated in Case 2.2).

Case 2.1: it is instructive to look at the special case of distance 1 from $\mathrm{tree}(\mathbf{w})$ and then argue how to iterate for larger distance. So let w_k be at distance 1 from $\mathrm{tree}(\mathbf{w})$ in the sense that $w_k \in R_a[u]$ for some u in $\mathrm{tree}(\mathbf{w})$ at depth d, but w_k not incident with an a-hyperedge in $\mathrm{tree}(\mathbf{w})$ (else in Case 2.2). Since $\mathbb{M}, u \sim^{\ell-d+1} \mathbb{M}', u'$ for $u' = \zeta(u)$ and

$$\mathfrak{M} \restriction [u]_a, \rho_{\sim^m}, u \equiv_r \mathfrak{M}' \restriction [u']_a, \rho_{\sim^m}, u' \quad \text{for } m = \ell - d + 1, \text{[16]}$$

a suitable response to the move that pebbles the world w_k (or the singleton information state $\{w_k\}$) in that game yields a world $w'_k \in E_a[u']$ such that

---

[16] Recall that $\sim^1$ above $\rho_{\sim^m}$ suffices for Corollary 3.21; and $\ell - d + 1 = m$.

$\mathbb{M}, w_k \sim^{\ell-d} \mathbb{M}', w_k'$ and

$$\mathfrak{M} \upharpoonright [u]_a, \rho_{\sim^m}, \{u\}, \{w_k\} \equiv_{r-1} \mathfrak{M}' \upharpoonright [u']_a, \rho_{\sim^m}, \{u'\}, \{w_k'\} \quad \text{for } m = \ell - d.$$

These levels of equivalence and granularity are appropriate since the depth of w_k and w_k' is $d+1$, and since one new vertex contributes to the new hyperedge. So we may extend the isomorphism ζ to map w_k to w_k' in keeping with conditions (1) and (2); in particular $w_k \sim^{\ell-(d+1)+1} w_k'$. If w_k is at greater distance from its nearest neighbour u in tree($\mathbf{w}$) we can iterate this process of introducing new hyperedges with one new element at a time, degrading the level of inquisitive bisimulation equivalence by 1 in every step that takes us one step further away from the root, but maintaining equivalences $\equiv_{r-1}$ between the newly added local a-structures (recall that $\sim^1$ supports full bisimilarity of local structures so that this feature does not degrade with distance, while granularity does).

Case 2.2: this is the case of $w_k \in [u]_a$ for some u in tree($\mathbf{w}$) that is already incident with an a-hyperedge of tree($\mathbf{w}$). By (2) we have

$$\mathfrak{M} \upharpoonright [u]_a, \rho_{\sim^m}, \mathbf{s} \equiv_{r-z+1} \mathfrak{M}' \upharpoonright [u']_a, \rho_{\sim^m}, \mathbf{s}' \quad \text{for } m = \ell - d + 1,$$

where z is the size of the tuples $\mathbf{s}$ and $\mathbf{s}'$ already incident with these local component structures, and d is the depth of u and u'. So we can find a response to a move on $\{w_k\}$ in this game that yields $w_k' \in [u']_a$ such that $\mathbb{M}, w_k \sim^{\ell-d} \mathbb{M}', w_k'$ and

$$\mathfrak{M} \upharpoonright [u]_a, \rho_{\sim^m}, \mathbf{s}\{w_k\} \equiv_{r-z} \mathfrak{M}' \upharpoonright [u']_a, \rho_{\sim^m}, \mathbf{s}'\{w_k'\} \quad \text{for } m = \ell - d.$$

The levels of bisimulation equivalence and granularity are appropriate as $m = \ell - (d+1) + 1$, the depth of w_k and w_k' is one greater than that of u and u', and as the length of the tuples $\mathbf{s}$ and $\mathbf{s}'$ has been increased by 1. So we may extend ζ by matching w_k with w_k' and extending tree($\mathbf{w}$) and tree($\mathbf{w}$)' by these new vertices and stay consistent with conditions (1) and (2).

Case 3. If $\mathbf{I}$ chooses some $(w_k)_a$ in the first sort of a local constituent $\mathfrak{M} \upharpoonright [w_k]_a$ of $\mathfrak{N}$, the choice of an appropriate match w_k' for w_k is carried out as in Case 2 (with a label $a_k := a$ in the description of the position), and $\mathbf{II}$ is to respond with a pebble on $(w_k')_a$. This finishes the proof of Lemma 3.22 in its essentially world-oriented formulation. $\square$

The adaptations for the non-trivially state-pointed version are marginal. Recall the illustration of the upgrading target in Figure 4, equation (6) and Proposition 3.8. By Definition 3.2 we use, for the state-pointed case, $\mathsf{I}(\mathbb{M}, s) = \mathsf{I}(\mathfrak{M}, s) = \mathsf{I}(\mathfrak{M}_s), s$, based on the augmentation $\mathfrak{M}_s$ of $\mathfrak{M} = \mathfrak{M}^{\mathsf{lf}}(\mathbb{M})$ by the dummy agent $s \in \mathcal{A}_s$ and parameter $s = s_{\alpha_s}$ in the second sort. So it suffices to establish

$$\mathfrak{N}, s := \mathsf{I}(\mathfrak{M}_s) \upharpoonright N^\ell(s), s \equiv_r^{(\ell)} \mathsf{I}(\mathfrak{M}_{s'}) \upharpoonright N^\ell(s'), s' =: \mathfrak{N}', s'$$

for suitable ℓ and r, where $\mathbb{M}_s, s \sim^n \mathbb{M}'_{s'}, s'$ for $n \geqslant \ell + 1$, and both $\mathbb{M}_s$ and $\mathbb{M}'_{s'}$ are N-acyclic, now for $N > 2(\ell + 2)$, sufficiently simple and rich (cf. Proposition 3.18). Here $\mathbb{M}, s \sim^n \mathbb{M}', s'$ implies that for any choice of $w \in s$ we can find $w' \in s'$ such that $\mathbb{M}, w \sim^n \mathbb{M}', w'$. We note that $N^\ell(s) \subseteq N^{\ell+2}(w)$ and that $s = s_{\alpha_s}$ as well as $N^\ell(s)$ are FO-definable in $I(\mathfrak{M}_s) \upharpoonright N^{\ell+1}(w), w$ (and similarly on the side of $\mathbb{M}'$). It therefore suffices to establish $\equiv_{r+t}^{(\ell+2)}$ between $\mathfrak{N}, w := I(\mathfrak{M}_s) \upharpoonright N^{\ell+2}(w), w$ and its counterpart $\mathfrak{N}', w' := I(\mathfrak{M}_{s'}) \upharpoonright N^{\ell+2}(w'), w'$ (for a constant t depending on just ℓ that is large enough to cover the FO-definitions of s and $N^\ell(s)$ in terms of $w \in s$ in the exploded views). Modulo these adaptations of the parameters, this case therefore reduces to the world-pointed case treated in Lemma 3.22. The initial condition $\mathsf{Inv}(0)$ now concerns, on the side of $\mathfrak{N}$ say, the anchor $w_0 := w$ in the core, augmented by $a_0 := s \in \mathcal{A}_s$ and $s_0 := s_{\alpha_s} = [w]_s$ as induced by the initial pebble on s_{α_s} in the second sort of the s-local component structure $\mathfrak{M}_s \upharpoonright [w]_s$ in the exploded view.

3.4.3 Summary: back to the main theorem

Returning to the main theorem, Theorem 1.2, we recall how Corollary 2.10 reduces the INQML-definability claim for $\sim$-invariant (and for state properties, inquisitive) FO-defined world- or state-properties over the respective classes of relational inquisitive epistemic models to their $\sim^n$-invariance, for some sufficiently large finite value for n; see Observation 3.1 to this effect. This criterion can be established through an upgrading of a sufficiently high level of $\sim^n$-equivalence to indistinguishability by the given FO-formula. Through a chain of $\sim$-preserving model transformations within the class of models under consideration, and by passage to the exploded view of these models, this upgrading eventually boils down to an upgrading of $\sim^n$ to some level of local Gaifman equivalence, as stated in Proposition 3.8. Figure 4 provides an overview over the several translation steps in the two tracks for the world-pointed and the state-pointed version, respectively. And those upgrading tracks have been completed with Lemma 3.22 and its adaptation to the non-trivially state-pointed case.

So this concludes the proof of the main theorem – in the world-pointed and in the state-pointed versions, respectively. In each case the argument is good in the classical as well as in the finite model theory reading since all the intermediate stages in the upgrading argument, within the class of general or locally full relational encodings, also preserve finiteness. Via passage to the exploded views, the upgrading always goes through locally full relational encodings. The difference between the expressive completeness assertions over the class of general versus locally full relational encodings may therefore almost be overlooked; but if, for instance $\varphi = \varphi(\mathsf{s}) \in$ FO is only assumed to be $\sim$-invariant and persistent over the class of locally full relational encodings, it can also only be guaranteed to be $\sim^n$-invariant for sufficiently large n over those encodings, and hence equivalently expressible by some $\varphi' \in$ INQML over that restricted class of relational encodings; it is the logical equivalence $\varphi' \equiv \varphi$ that is subject to the restriction; and of course $\varphi' \equiv \varphi$ remains valid over the class

of all relational encodings provided φ is $\sim$-invariant throughout, just as φ' is anyway.

4   Conclusion

In our previous paper [10] we introduced a notion of bisimulation for inquisitive modal logic and proved a counterpart of the van Benthem–Rosen theorem for inquisitive modal logic over the class of all (finite or arbitrary) relational encodings of inquisitive modal models. In this paper, we have extended those results to the especially interesting, but technically challenging, case of inquisitive *epistemic* models, which satisfy $S5$-like constraints. We have compared the expressive power of inquisitive modal logic to that of first-order logic interpreted over relational encodings of such models, showing that a property of world-pointed inquisitive epistemic models is expressible in inquisitive epistemic logic just in case it is expressible in first-order logic and bisimulation invariant; moreover, the same goes for *inquisitive* properties of state-pointed models, i.e., properties that are preserved in subsets and always contain the empty state. We have seen that the result holds also if we restrict to the class of finite models; moreover, it holds regardless of whether or not we impose a *local fullness* constraint, which requires the encoding of a model to represent in the second sort the full powerset of each equivalence class in the underlying Kripke frame.

Conceptually, as we discussed in the introduction, the result is interesting as it implies that inquisitive modal logic is in a precise sense *expressively complete* as a language to specify bisimulation-invariant properties of worlds in inquisitive epistemic models (or bisimulation-invariant inquisitive properties of information states). Technically, we deal with first-order definability over non-elementary classes of 2-sorted relational structures that give some specifically constrained access to sets of subsets in their second sort. This takes us into explorations of a border area between first-order and monadic second-order expressiveness. As it turns out, the model classes in question are still sufficiently malleable up to bisimulation to allow for suitably adapted locality arguments in a back-and-forth analysis – albeit after having translated the structures in question into a more manageable regime by means of first-order interpretations. In this context an adaptation of well-established classical model-theoretic methods is at the core of our treatment; and the specific combination of these ingredients can here serve rather intricate non-classical ends. It is in this respect that, also technically, the present paper goes substantially beyond [10].

Several directions for further work naturally suggest themselves. First, we may hope to extend this result to other inquisitive modal logics determined by salient classes of frames; for instance, to the inquisitive counterpart of the modal logics K4 (determined by the condition: if $v \in \sigma_a(w)$ then $\Sigma_a(v) \subseteq \Sigma_a(w)$, which implies the transitivity of the underlying Kripke frame) and S4 (determined by the above condition plus the requirement $w \in \sigma_a(w)$, which amounts to the reflexivity of the underlying Kripke frame). Characterisations of bisimulation-invariant FO over related classes of Kripke structures have been obtained in [14];

for discussion of the inquisitive counterparts of some standard frame conditions, see §7.4 of [4].

Moreover, recent work [3] introduced a generalization of inquisitive modal logic designed to describe models where the set $\Sigma(w)$ of successors of a given world w is not necessarily downward closed. This generalization uses a conditional modality $\Rrightarrow$, where a formula $\varphi \Rrightarrow \psi$ is interpreted as claiming that every successor state that supports φ also supports ψ. One interesting question is whether an analogue of our main theorem holds for this generalised inquisitive logic relative to the class of all $S5$-like models, i.e., models where we have $w \in \sigma_a(w)$ for all w, and moreover $v \in \sigma_a(w)$ implies $\Sigma_a(w) = \Sigma_a(v)$, but where the set of successors $\Sigma(w)$ is not necessarily downward closed.

References

[1] S. Abramsky and J. Väänänen. From IF to BI. *Synthese*, 167(2):207–230, 2009.

[2] I. Ciardelli. Modalities in the realm of questions: axiomatizing inquisitive epistemic logic. In R. Goré, B. Kooi, and A. Kurucz, editors, *Advances in Modal Logic (AIML)*, pages 94–113, London, 2014. College Publications.

[3] I. Ciardelli. Describing neighborhoods in inquisitive modal logic. In A. Fernandez-Duque, A. Palmigiano, and S. Pinchinay, editors, *Advances in Modal Logic (AIML)*, London, 2022. College Publications.

[4] I. Ciardelli. *Questions in Logic*, volume DS-2016-01 of *ILLC Dissertation Series*. Institute for Logic, Language and Computation, Amsterdam, 2015.

[5] I. Ciardelli. Dependency as question entailment. In S. Abramsky, J. Kontinen, J. Väänänen, and H. Vollmer, editors, *Dependence Logic: theory and applications*, pages 129–181. Springer International Publishing Switzerland, 2016.

[6] I. Ciardelli. Questions as information types. *Synthese*, 195(1):321–365, 2018.

[7] I. Ciardelli, J. Groenendijk, and F. Roelofsen. Inquisitive semantics: A new notion of meaning. *Language and Linguistics Compass*, 7(9):459–476, 2013.

[8] I. Ciardelli, J. Groenendijk, and F. Roelofsen. On the semantics and logic of declaratives and interrogatives. *Synthese*, 192(6):1689–1728, 2015.

[9] I. Ciardelli and M. Otto. Bisimulation in inquisitive modal logic. In Jerôme Lang, editor, *Theoretical Aspects of Rationality and Knowledge (TARK) 16*. EPTCS, 2017.

[10] I. Ciardelli and M. Otto. Inquisitive bisimulation. *Journal of Symbolic Logic*, 86(1):77–109, 2021.

[11] I. Ciardelli and F. Roelofsen. Inquisitive logic. *Journal of Philosophical Logic*, 40:55–94, 2011.

[12] I. Ciardelli and F. Roelofsen. Inquisitive dynamic epistemic logic. *Synthese*, 192(6):1643–1687, 2015.

[13] I. Ciardelli. Inquisitive Logic. Consequence and inference in the realm of questions. *Springer*, 2023.

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

[15] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Perspectives in Mathematical Logic. Springer, 2 edition, 1999.

[16] H. Gaifman. On local and nonlocal properties. In J. Stern, editor, *Logic Colloquium 81*, pages 105–135. North Holland, 1982.

[17] P. Galliani. Inclusion and exclusion dependencies in team semantics – on some logics of imperfect information. *Annals of Pure and Applied Logic*, 163(1):68–84, 2012.

[18] T. van Gessel. Action models in inquisitive logic. *Synthese*, 197:3905–3945, 2020.

[19] V. Goranko and M. Otto. Model theory of modal logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, pages 249–329. Elsevier, 2007.

[20] H. H. Hansen. Monotone modal logics. MSc Thesis, University of Amsterdam, 2003.

[21] H. H. Hansen, C. Kupke, and E. Pacuit. Neighbourhood structures: Bisimilarity and basic model theory. *Logical Methods in Computer Science*, 2, 2009.

[22] J. Hintikka. *Knowledge and Belief: an Introduction to the Logic of the Two Notions*. Cornell University Press, 1962.

[23] W. Hodges. Compositional semantics for a language of imperfect information. *Logic Journal of IGPL*, 5(4):539–563, 1997.

[24] W. Hodges. Some strange quantifiers. In *Structures in logic and computer science. Lecture notes in computer science*, pages 51–65. Springer, New York, 1997.

[25] M. Otto. Elementary proof of the van Benthem–Rosen characterisation theorem. Technical Report 2342, Fachbereich Mathematik, Technische Universität Darmstadt, 2004.

[26] M. Otto. Modal and guarded characterisation theorems over finite transition systems. *Annals of Pure and Applied Logic*, 130:173–205, 2004.

[27] M. Otto. Highly acyclic groups, hypergraph covers and the guarded fragment. *Journal of the ACM*, 59 (1), 2012.

[28] S. Meissner and M. Otto. A first-order framework for inquisitive modal logic. *The Review of Symbolic Logic*, 15(2):311-333, 2022.

[29] E. Pacuit. *Neighborhood Semantics for Modal Logic*. Springer, 2017. forthcoming.

[30] F. Roelofsen. Algebraic foundations for the semantic treatment of inquisitive content. *Synthese*, 190(1):79–102, 2013.

[31] J. Plaza. Logics of public communications. *Synthese*, 158:165–179, 2007.

[32] V. Punvcochávr and I. Sedlár.
Epistemic extensions of substructural inquisitive logics. *Journal of Logic and Computation*, 31 (7):1820–1844, 2021.

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

[34] J. Väänänen. *Dependence Logic: A New Approach to Independence Friendly Logic*. Cambridge University Press, 2007.

[35] J. Väänänen. Modal dependence logic. In K. R. Apt and R. van Rooij, editors, *New Perspectives on Games and Interaction*, pages 237–254. Amsterdam University Press, 2008.

[36] J. van Benthem. *Modal Logic and Classical Logic*. Bibliopolis, Napoli, 1983.

[37] F. Yang. *On extensions and variants of dependence logic: A study of intuitionistic connectives in the team semantics setting*. PhD thesis, University of Helsinki, 2014.

[38] F. Yang and J. Väänänen. Propositional logics of dependence. *Annals of Pure and Applied Logic*, 167:557–589, 2016.