Hierarchies in independence logic*

Pietro Galliani, Miika Hannula, and Juha Kontinen

University of Helsinki, Department of Mathematics and Statistics,
P.O. Box 68, 00014 Helsinki, Finland
`pgallian@gmail.com, {miika.hannula,juha.kontinen}@helsinki.fi`

Abstract

We study the expressive power of fragments of inclusion and independence logic defined either by restricting the number of universal quantifiers or the arity of inclusion and independence atoms in formulas. Assuming the so-called lax semantics for these logics, we relate these fragments of inclusion and independence logic to familiar sublogics of existential second-order logic. We also show that, with respect to the stronger strict semantics, inclusion logic is equivalent to existential second-order logic.

1 Introduction

Independence logic [15] and inclusion logic [11] are recent variants of dependence logic. Dependence logic [20] extends first-order logic by dependence atomic formulas

$$=(x_1, \ldots, x_n) \tag{1}$$

the meaning of which is that the value of x_n is completely determined by the values of $x_1, \ldots, x_{n-1}$. The semantics of dependence logic is defined using sets of assignments rather than a single assignment as in first-order logic. Independence logic replaces the dependence atoms by independence atoms $\vec{y} \perp_{\vec{x}} \vec{z}$, the intuitive meaning of which is that, with respect to any fixed value of $\vec{x}$, the variables $\vec{y}$ are totally independent of the variables $\vec{z}$. In inclusion logic dependence atoms are replaced by inclusion atoms $\vec{x} \subseteq \vec{y}$, meaning that all the values of $\vec{x}$ appear also as values for $\vec{y}$. We study the expressive power of the syntactic fragments of these logics defined either by restricting the number of universal quantifiers or the arity of the independence and inclusion atoms in sentences. These results are proved with respect to lax semantics. We also show that, under strict semantics, inclusion logic is equivalent to existential second-order logic ESO while, by a recent result of Hella and Galliani [3], with lax semantics inclusion logic is equivalent to greatest fixed point logic, and hence to LFP (and PTIME) over finite (ordered) structures.

Since the introduction of dependence logic ($\mathcal{D}$) in 2007 many interesting variants of it have been introduced. In fact the team semantics of dependence logic has turned into a general framework for logics in which various notions of dependence and independence can be formalized. Dependence logic has a very intimate and well understood connection to ESO dating back to the results of [17, 8, 22] on Henkin quantifiers. For some of the new variants

and concepts in this area the correspondence to ESO does not hold. We briefly mention some related work on the complexity theoretic aspects of these logics:

- The extension of dependence logic by so-called intuitionistic implication $\rightarrow$ (introduced in [1]) increases the expressive power of dependence logic to full second-order logic [23].
- The model checking problem of full dependence logic, and many of its variants, was recently shown to be NEXPTIME-complete. In fact, for any variant of dependence logic whose atoms are PTIME-computable, the corresponding model checking problem is contained in NEXPTIME [13].
- The non-classical interpretation of disjunction in dependence logic has the effect that the model checking problems of $\phi_1 := \; =(x,y) \lor \; =(u,v)$ and $\phi_2 := \; =(x,y) \lor \; =(u,v) \lor \; =(u,v)$ are already NL-complete and NP-complete, respectively [18].

While dependence logic and independence logic are both equivalent to ESO in expressive power [20, 15], for inclusion logic only containment in ESO has been shown [11]. Furthermore, the expressive power of various natural syntactic fragments of independence and inclusion logics is not understood at the moment. The starting point of our work were the results of [7] on the fragments $\mathcal{D}(k\forall)$ and $\mathcal{D}(k\text{-dep})$ of dependence logic. The fragment $\mathcal{D}(k\forall)$ contains those $\mathcal{D}$-formulas in which at most k variables have been universally quantified, and in the formulas of $\mathcal{D}(k\text{-dep})$ dependence atoms of arity at most k may appear (atoms of the form $=(x_1, \ldots, x_n)$ satisfying $n \leq k+1$). The following results were shown in [7]:

1. $\mathcal{D}(k\text{-dep}) = \text{ESO}_f(k\text{-ary})$,
2. $\mathcal{D}(k\forall) \leq \text{ESO}_f(k\forall) \leq \mathcal{D}(2k\forall)$

where $\text{ESO}_f(k\text{-ary})$ is the fragment of ESO in which the quantified functions and relations have arity at most k, and $\text{ESO}_f(k\forall)$ consists of ESO-sentences that are in Skolem Normal Form and contain at most k universal first-order quantifiers. The equivalence in (1) was used to show that in $\mathcal{D}(k\text{-dep})$ even cardinality of a $k+1$-ary relation cannot be expressed using the result of Ajtai [2]. On the other hand, since

$$\text{ESO}_f(k\forall) = \text{NTIME}_{\text{RAM}}(n^k) < \text{NTIME}_{\text{RAM}}(n^{k+1})$$

by [14] and [6], an infinite expressivity hierarchy for the fragments $\mathcal{D}(k\forall)$ was shown using 2. Above $\text{NTIME}_{\text{RAM}}(n^k)$ denotes the family of classes of τ-structures that can be recognized by a non-deterministic RAM in time $O(n^k)$.

In [11] it was observed that independence logic and inclusion logic can be given two alternative semantics called strict and lax semantics. For dependence logic these two semantics coincide in the sense that the meaning of any $\mathcal{D}$-formula is the same under both interpretations. For independence and inclusion logic formulas this is not the case as shown in [11]. In fact, we will show that, with respect to strict semantics, inclusion logic is equivalent to ESO, while by a recent result of Hella and Galliani [3], with lax semantics inclusion logic is equivalent to greatest fixed point logic. In the rest of the article we consider the expressive power of fragments of independence logic and inclusion logic with lax semantics. First we look at fragments defined analogously to $\mathcal{D}(k\text{-dep})$ of dependence logic. We let $\text{FO}(\perp_{\text{c}})(k\text{-ind})$ contain those independence logic sentences in which independence atoms with at most $k+1$ different variables may appear. Similarly in the sentences of $\text{FO}(\subseteq)(k\text{-inc})$ only inclusion atoms of the form $\vec{a} \subseteq \vec{b}$, where $|\vec{a}| = |\vec{b}| \leq k$ may appear. Our results show that

$$\text{FO}(\subseteq)(k\text{-inc}) \leq \text{ESO}_f(k\text{-ary}) = \text{FO}(\perp_{\text{c}})(k\text{-ind}).$$

Then we consider the analogues of $\mathcal{D}(k\forall)$ in the case of $\text{FO}(\perp) = \text{FO}(\perp_{\text{c}})$ [21], which is the sublogic of independence logic allowing only so-called pure atoms $\vec{y} \perp \vec{z}$, and $\text{FO}(\perp, \subseteq)$. We show that

- $FO(\bot)(2\forall) = FO(\bot)$,
- $FO(\bot, \subseteq)(1\forall) = FO(\bot, \subseteq)$.

This article is organized as follows. In Section 2 we review some basic properties and results regarding dependence logic and its variants. In Section 3 we compare the strict and lax semantics and in Section 4.1 relate the arity fragments of independence logic and inclusion logic with that of ESO. Finally, in Section 4.2 we consider fragments defined by restricting the number of universally quantified variables.

2 Preliminaries

2.1 Team Semantics

Team semantics is a generalization of Tarski semantics in which formulas are interpreted by *sets* of assignments, called *teams*, rather than by single assignments. In this subsection, we will recall the definition of team semantics for first order logic. We will assume that all our formulas are in negation normal form. Also, all structures considered in the paper are assumed to have at least two elements.

▶ **Definition 2.** *Let $\mathcal{M}$ be a first-order model and V a finite set of variables. Then*
- *a team X over $\mathcal{M}$ with domain $Dom(X) = V$ is a finite set of assignments from V to the domain M of $\mathcal{M}$;*
- *for a tuple $\vec{v}$ of variables in V, we write $X(\vec{v})$ for the set $\{s(\vec{v}) : s \in X\}$ of all values that $\vec{v}$ takes in X, where $s(\vec{v}) := (s(v_1), \ldots, s(v_n))$;*
- *for a subset W of V, we write $X \restriction W$ for the team obtained by restricting all assignments of X to the variables in W.*
- *For a formula ϕ, the set of free variables of ϕ is denoted by $\mathrm{Fr}(\phi)$.*

There exist two variants of team semantics, called respectively *strict* and *lax*, which differ with respect to the interpretation of disjunction and existential quantification. Informally speaking, the choice between strict and lax semantics corresponds to the choice between disallowing or allowing *nondeterministic strategies* in the corresponding semantic games.[1]

We first give the definition of the lax version of team semantics; later, we will discuss some of the ways in which strict semantics differs from it.

▶ **Definition 3** (Team Semantics). *Let $\mathcal{M}$ be any first-order model and let X be any team over it. Then*

TS-lit: *For all first-order literals α, $\mathcal{M} \models_X \alpha$ if and only if, for all $s \in X$, $\mathcal{M} \models_s \alpha$ in the usual Tarski semantics sense;*

TS-$\vee$: *For all ψ and θ, $\mathcal{M} \models_X \psi \vee \theta$ if and only if $X = Y \cup Z$ for two subteams Y and Z such that $\mathcal{M} \models_Y \psi$ and $\mathcal{M} \models_Z \theta$;*

TS-$\wedge$: *For all ψ and θ, $\mathcal{M} \models_X \psi \wedge \theta$ if and only if $\mathcal{M} \models_X \psi$ and $\mathcal{M} \models_X \theta$;*

TS-$\exists$: *For all ψ and all variables v, $\mathcal{M} \models_X \exists v \psi$ if and only if there exists a function $H : X \to \mathcal{P}(M)\backslash\{\emptyset\}$ such that $\mathcal{M} \models_{X[H/v]} \psi$, where $X[H/v] = \{s[m/v] : s \in X, m \in H(s)\}$;*

TS-$\forall$: *For all ψ and all variables v, $\mathcal{M} \models_X \forall v \psi$ if and only if $\mathcal{M} \models_{X[M/v]} \psi$, where $X[M/v] = \{s[m/v] : s \in X, m \in M\}$.*

---

[1] See [10] and [13] for details.

If $\mathcal{M} \models_X \phi$, we say that X satisfies ϕ in $\mathcal{M}$; and if a sentence (that is, a formula with no free variables) ϕ is satisfied by the team $\{\emptyset\}$,[2] we say that ϕ is true *in $\mathcal{M}$ and we write $\mathcal{M} \models \phi$.*

In the team semantics setting, formulas ϕ and ψ are said to be logically equivalent, $\phi \equiv \psi$, if for all models $\mathcal{M}$ and teams X, with $\mathrm{Fr}(\phi) \cup \mathrm{Fr}(\psi) \subseteq \mathrm{Dom}(X)$, $\mathcal{M} \models_X \phi \Leftrightarrow \mathcal{M} \models_X \psi$. Logics $\mathcal{L}$ and $\mathcal{L}'$ are said to be equivalent, $\mathcal{L} = \mathcal{L}'$, if every $\mathcal{L}$-sentence ϕ is equivalent to some $\mathcal{L}'$-sentence ψ, and vice versa.

The following result can be proved by structural induction on the formula ϕ:

▶ **Theorem 4** (Flatness). *For all first order formulas ϕ and all suitable models $\mathcal{M}$ and teams X, the following are equivalent:*

1. *$\mathcal{M} \models_X \phi$;*
2. *For all $s \in X$, $\mathcal{M} \models_{\{s\}} \phi$;*
3. *For all $s \in X$, $\mathcal{M} \models_s \phi$ according to Tarski semantics.*

2.2   Dependencies in Team Semantics

The advantage of team semantics, and the reason for its development, is that it allows us to extend first-order logic by new atoms and operators. For the purposes of this paper, the following atoms, inspired by database-theoretic *dependency notions*[3], are of particular interest:

▶ **Definition 5.** ▬ *Let $\vec{x}$ be a tuple of variables and let y be another variable. Then $=(\vec{x}, y)$ is a* dependence atom, *with the semantic rule*

   TS-dep: *$\mathcal{M} \models_X =(\vec{x}, y)$ if and only if any two $s, s' \in X$ which assign the same value to $\vec{x}$ also assign the same value to y;*

▬ *Let $\vec{x}$, $\vec{y}$, and $\vec{z}$ be tuples of variables (not necessarily of the same length). Then $\vec{y} \perp_{\vec{x}} \vec{z}$ is a* conditional independence atom, *with the semantic rule*

   TS-ind: *$\mathcal{M} \models_X \vec{y} \perp_{\vec{x}} \vec{z}$ if and only if for any two $s, s' \in X$ which assign the same value to $\vec{x}$ there exists a $s'' \in X$ which agrees with s with respect to $\vec{x}$ and $\vec{y}$ and with s' with respect to $\vec{z}$.*

   Furthermore, we will write $\vec{x} \perp \vec{y}$ as a shorthand for $\vec{x} \perp_\emptyset \vec{y}$, and call it a pure independence atom*;*

▬ *Let $\vec{x}$ and $\vec{y}$ be two tuples of variables of the same length. Then $\vec{x} \subseteq \vec{y}$ is an* inclusion atom, *with the semantic rule*

   TS-inc: *$\mathcal{M} \models_X \vec{x} \subseteq \vec{y}$ if and only if $X(\vec{x}) \subseteq X(\vec{y})$;*

Given a collection $\mathcal{C} \subseteq \{=(\ldots), \perp_c, \subseteq\}$ of atoms, we will write $\mathrm{FO}(\mathcal{C})$ (omitting the set parenthesis of $\mathcal{C}$) for the logic obtained by adding them to the language of first-order logic. With this notation dependence logic, independence logic and inclusion logic are denoted by $\mathrm{FO}(=(\ldots))$, $\mathrm{FO}(\perp_c)$ and $\mathrm{FO}(\subseteq)$, respectively. We will also write $\mathrm{FO}(\perp)$ for the fragment of independence logic containing only pure independence atoms.

---

[2] $\{\emptyset\}$ is the team containing the empty assignment. Of course, this is different from the *empty team* $\emptyset$, containing no assignments.

[3] More precisely, dependence atoms correspond to functional dependencies [4], independence atoms to embedded multivalued dependencies and conditional dependency conditions as in [12, 19], and inclusion atoms to inclusion dependencies [9, 5].

All formulas of all the above-mentioned logics satisfy the two following properties:

▶ **Proposition 6** (Empty Team Property). *For all models $\mathcal{M}$ and $\phi \in \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$ over the signature of $\mathcal{M}$, $\mathcal{M} \models_\emptyset \phi$.*

▶ **Proposition 7** (Locality). *Let ϕ be a formula of $\mathrm{FO}(=(\ldots), \perp_c, \subseteq)$ whose free variables $\mathrm{Fr}(\phi)$ are contained in V. Then, for all models $\mathcal{M}$ and teams X, $\mathcal{M} \models_X \phi$ if and only if $\mathcal{M} \models_{X \restriction V} \phi$.*

Furthermore, we have the two following results for dependence logic:

▶ **Proposition 8** (Downwards Closure). *For all models $\mathcal{M}$, dependence logic formulas ϕ and teams X, if $\mathcal{M} \models_X \phi$ then $\mathcal{M} \models_Y \phi$ for all $Y \subseteq X$.*

▶ **Theorem 9** ([22, 8, 20]). *Any dependence logic sentence ϕ is logically equivalent to some ESO sentence ϕ^*, and vice versa.*

What about independence logic? As shown in [15], a dependence atom $=(\vec{x}, y)$ is logically equivalent to the independence atom $y \perp_{\vec{x}} y$, and, since independence logic is clearly contained in ESO, we have at once that

▶ **Theorem 10** ([15]). *Any independence logic sentence ϕ is logically equivalent to some ESO sentence ϕ^*, and vice versa.*

Furthermore,

▶ **Theorem 11** ([21]). *Any independence logic formula is equivalent to some pure independence logic formula.*

For inclusion logic the following is known.

▶ **Theorem 12.**
1. *An inclusion atom $\vec{x} \subseteq \vec{y}$ is equivalent to the $\mathrm{FO}(\perp)$ expression*

$$\forall v_1 v_2 \vec{z}((\vec{z} \neq \vec{x} \wedge \vec{z} \neq \vec{x}) \vee (v_1 \neq v_2 \wedge \vec{z} \neq \vec{y}) \vee ((v_1 = v_2 \vee \vec{z} = \vec{y}) \wedge \vec{z} \perp v_1 v_2))$$

where v_1, v_2 and $\vec{z}$ are new variables [11].
2. *Any inclusion logic sentence ϕ is logically equivalent to some positive greatest fixpoint logic sentence ϕ^*, and vice versa [3].*

We conclude this subsection with two novel results, a characterization of *dependence* in terms of pure independence and a *prenex normal form theorem* for formulas of our logics.

▶ **Theorem 13.** *For all models $\mathcal{M}$ and teams X*

$$\mathcal{M} \models_X =(\vec{x}, y) \Leftrightarrow \mathcal{M} \models_X \forall \vec{z} \exists w((\vec{z} = \vec{x} \rightarrow w = y) \wedge \vec{x}y \perp \vec{z}w).$$

▶ **Lemma 14.** *Let $\phi, \psi \in \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$ and let x be a variable not occurring free in ψ. Then the following equivalences hold:*
1. $\exists x \phi \wedge \psi \equiv \exists x(\phi \wedge \psi)$,
2. $\exists x \phi \vee \psi \equiv \exists x(\phi \vee \psi)$,
3. $\forall x \phi \wedge \psi \equiv \forall x(\phi \wedge \psi)$,
4. $\forall x \phi \vee \psi \equiv \exists a \exists b \forall x((\phi \wedge a = b) \vee (\psi \wedge a \neq b))$ *where a and b are new variables.*

Lemma 14 allows us to show the following.

▶ **Theorem 15.** *Any formula $\phi \in \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$ is logically equivalent to some formula ϕ' such that*
1. ϕ' *is of the form $Q_1 x_1 \ldots Q_k x_k \psi$, where ψ is quantifier-free;*
2. *Any literal or non-first-order atom which occurs in ϕ' occurred already in ϕ;*
3. *The number of universal quantifiers in ϕ' is the same as the number of universal quantifiers in ϕ.*

3 Comparing strict and lax semantics

As we mentioned, there exists an alternative variant of lax semantics, called strict semantics. It differs from lax semantics in the definition of the semantic rules for disjunction and existential quantification, which are replaced respectively by

STS-$\vee$**:** For all ψ and θ, $\mathcal{M} \models_X \psi \vee \theta$ if and only if Y and Z exist such that $Y \cup Z = X$, $Y \cap Z = \emptyset$, $\mathcal{M} \models_Y \psi$ and $\mathcal{M} \models_Z \theta$;

STS-$\exists$**:** For all ψ and all variables v, $\mathcal{M} \models_X \exists v \psi$ if and only if there exists a function $F : X \to M$ such that $\mathcal{M} \models_{X[F/v]} \psi$, where $X[F/v] = \{s[F(s)/v] : s \in X\}$.

In the original version of dependence logic lax disjunction and strict existential quantification were used [20]. However, since dependence logic is downwards closed, it does not make any difference whether strict or lax version of disjunction (or existential quantification) is used. In general the following holds.

▶ **Proposition 16.** *If* $\mathcal{M} \models_X \phi$ *according to strict team semantics, then* $\mathcal{M} \models_X \phi$ *according to lax team semantics.*

For downwards closed logics, such as dependence logic, the converse is then also true.

▶ **Proposition 17** ([11])**.** *For all dependence logic formulas* ϕ*, models* $\mathcal{M}$ *and teams* X*,* $\mathcal{M} \models_X \phi$ *holds wrt strict team semantics if and only if it holds wrt lax team semantics.*

However, the same is false for both inclusion logic and independence logic. In particular, as we will now see, inclusion logic with strict semantics is equivalent to full existential second order logic, in contrast with the second item of Theorem 12.

By Theorem 9, it suffices to show that every dependence logic sentence is equivalent to some inclusion logic sentence (with strict semantics). In order to do so, we will use the following *normal form theorem* from [20]:

▶ **Theorem 18** ([20])**.** *Every dependence logic sentence is equivalent to some sentence of the form*

$$\phi := \forall \vec{x} \exists \vec{y} \left(\bigwedge_{y_i \in \vec{y}} =(\vec{v_i}, \vec{y_i}) \wedge \theta \right) \tag{19}$$

where for all i*,* $\vec{v_i}$ *is contained in* $\vec{x}$ *and where* θ *is a quantifier-free first-order formula.*

As we will now show, in strict semantics the dependence atoms in (19) can be replaced by equivalent inclusion logic subformulas; and, therefore, it follows at once that (strict) inclusion logic is equivalent to dependence logic (and, therefore, to ESO) over sentences.

▶ **Definition 20.** *Let* $\mathcal{M}$ *be a model and* X *a team, and let* $\vec{x}$ *be a tuple of variables in its domain. We say that* X *is* $\vec{x}$*-universal if for all tuples of elements* $\vec{m}$ *with* $|\vec{m}| = |\vec{x}|$*, there exists one and only one* $s \in X$ *with* $s(\vec{x}) = \vec{m}$*.*

▶ **Lemma 21.** *If* X *is of the form* $\{\emptyset\}[M/\vec{x}][\vec{F}/\vec{y}]$ *then* X *is* $\vec{x}$*-universal.*

Proof. Obvious (but note that if the $\vec{F}$ were replaced by *nondeterministic* choice functions $\vec{H}$, as in the case of the lax semantics, this would not hold). ◀

▶ **Proposition 22.** *Let* $\mathcal{M}$ *be a model and* X *a* $\vec{x}$*-universal team. Suppose also that* $y \notin \vec{x}$*,* $\vec{v} \subseteq \vec{x}$*, and* $\vec{w} = \vec{x} \backslash \vec{v}$ *(that is,* $\vec{w}$ *lists, without repetitions, all variables occurring in* $\vec{x}$ *but not in* $\vec{v}$*). Then*

$$\mathcal{M} \models_X =(\vec{v}, y) \Leftrightarrow \mathcal{M} \models_X \forall \vec{q} (\vec{q}\vec{v}y \subseteq \vec{w}\vec{v}y).$$

Proof. Suppose that $\mathcal{M} \models_X = (\vec{v}, y)$, and let $h = s[\vec{m}/\vec{q}] \in X[M/\vec{q}]$, where $s \in X$. Since X is $\vec{x}$-universal and $\vec{x} = \vec{v} \cup \vec{w}$, there exists an assignment $s' \in X$ such that $s'(\vec{w}) = \vec{m}$ and $s'(\vec{v}) = s(\vec{v})$. Since y is a function of $\vec{v}$ alone, this implies that $s'(y) = s(y)$. Finally, $h' = s'[\vec{m}/\vec{q}] \in X[M/\vec{q}]$, and $h'(\vec{w}\vec{v}y) = \vec{m}s(\vec{v}y) = h(\vec{q}\vec{v}y)$, as required.

Conversely, suppose that $\mathcal{M} \models_X \forall \vec{q}(\vec{q}\vec{v}y \subseteq \vec{w}\vec{v}y)$, and let $s, s' \in X$ be such that $s(\vec{v}) = s'(\vec{v})$. Now let $\vec{m} = s'(\vec{w})$, and consider $h = s[\vec{m}/\vec{q}] \in X[M/\vec{q}]$. By hypothesis, there exists a $h' \in X[M/\vec{q}]$ such that $h'(\vec{w}) = h(\vec{q}) = \vec{m}$ and $h'(\vec{v}y) = h(\vec{v}y) = s(\vec{v}y)$. This h' is of the form $s''[\vec{m'}/\vec{q}]$ for some $s'' \in X$; and for this s'', we have that $s''(\vec{v}) = s(\vec{v}) = s'(\vec{v})$, $s''(\vec{w}) = \vec{m} = s'(\vec{w})$ and $s''(y) = s(\vec{y})$. Now, $\vec{x} = \vec{v} \cup \vec{w}$, and s'' coincides with s' over it, and X is $\vec{x}$-universal; therefore, we have to conclude that $s'' = s'$. But then $s'(y) = s''(y) = s(y)$, and therefore s'' and s coincide over y too. ◄

▶ **Corollary 23.** *With strict semantics inclusion logic is equivalent to* ESO.

Proof. By Lemma 21 and the Proposition 22, any sentence of the form (19) can be expressed in inclusion logic as

$$\forall \vec{x} \exists \vec{y} \left(\bigwedge_{y_i \in \vec{y}} (\forall \vec{q_i}(\vec{q_i}\vec{v_i}y \subseteq \vec{w_i}\vec{v_i}y)) \wedge \theta \right) \tag{24}$$

where for all i, $\vec{w_i} = \vec{x} \backslash \vec{v_i}$; and this implies our result. ◄

The analogue of Theorem 7 (locality) for inclusion logic with strict semantics fails. As an especially surprising example of such an failure we now show that one can find inclusion logic sentences that count the number of assignments in a team:

▶ **Theorem 25.** *For each natural number n there is a sentence $\phi \in \mathrm{FO}(\subseteq)$ such that for all models $\mathcal{M}$ and teams X where $X \neq \emptyset$ and the variables in $Dom(X)$ do not appear in ϕ,*

$$\mathcal{M} \models_X \phi \text{ if and only if } |X| \geq n.$$

The failure of locality in non-downwards closed logics with strict semantics is somewhat problematic, as it causes the interpretation of a formula to depend on the values that our assignments take on variables which do *not* occur in it. As a consequence, in the rest of this work we will focus on logics with lax semantics.

4 The expressive power of fragments

The purpose of this section is to generalize the classification of the expressive power of fragments of dependence logic of [7] to the case of other variants (with respect to lax semantics). We will consider the following fragments.

▶ **Definition 26.** *Let $\mathcal{C}$ be a subset of $\{=(\ldots), \perp_c, \perp, \subseteq\}$ and let $k \in \mathbb{N}$. Then*
1. $\mathrm{FO}(\mathcal{C})(k{-}dep)$ *is the class of sentences of $\mathrm{FO}(\mathcal{C})$ in which dependence atoms of the form $=(\vec{z}, y)$, where $\vec{z}$ is of length at most k, may appear.*
2. $\mathrm{FO}(\mathcal{C})(k{-}ind)$ *is the class of sentences of $\mathrm{FO}(\mathcal{C})$ in which independence atoms of the form $\vec{y} \perp_{\vec{x}} \vec{z}$, where $\vec{x}\vec{y}\vec{z}$ has at most $k + 1$ distinct variables, may appear.*
3. $\mathrm{FO}(\mathcal{C})(k{-}inc)$ *is the class of sentences of $\mathrm{FO}(\mathcal{C})$ in which inclusion atoms of the form $\vec{a} \subseteq \vec{b}$, where $\vec{a}$ and $\vec{b}$ are of length at most k, may appear.*
4. $\mathrm{FO}(\mathcal{C})(k\forall)$ *is the class of sentences of $\mathrm{FO}(\mathcal{C})$ in which at most k universal quantifiers occur.*

As in [7], we will write $\mathcal{D}(k\text{-dep})$ and $\mathcal{D}(k\forall)$ for $\mathrm{FO}(=(\ldots))(k{-}dep)$ and $\mathrm{FO}(=(\ldots))(k\forall)$, respectively.

4.1   Arity hierarchies

In this section we will prove that $\mathrm{FO}(\perp_c)(k\text{-ind}) = \mathrm{ESO}_f(k\text{-ary})$. In particular this also implies that $\mathrm{FO}(\perp_c)(k\text{-ind}) = \mathcal{D}(k\text{-dep})$ [7]. We will also prove that $\mathrm{FO}(\subseteq)(k\text{-inc}) \leq \mathrm{ESO}_f(k\text{-ary})$. The direction from $\mathrm{ESO}_f(k\text{-ary})$ to $\mathrm{FO}(\perp_c)(k\text{-ind})$ is straightforward.

▶ **Proposition 27.** $\mathrm{ESO}_f(k\text{-ary}) \leq \mathrm{FO}(\perp_c)(k\text{-ind})$.

Proof. Let $\phi \in \mathrm{ESO}_f(k\text{-ary})$. By [7] there exists a $\phi' \in \mathcal{D}(k\text{-dep})$ equivalent to ϕ and of the form

$$Q^1 x_1 \ldots Q^m x_m \exists y_1 \ldots \exists y_n (\bigwedge_{1 \leq j \leq n} =(\vec{z}_j, y_j) \wedge \theta)$$

where $\vec{z}_j$, for $1 \leq j \leq n$, is a sequence of length at most k. By [15] each dependence atom $=(\vec{z}, y)$ is equivalent to the independence atom $y \perp_{\vec{z}} y$. Therefore we can present ϕ' in the following independence logic form

$$Q^1 x_1 \ldots Q^m x_m \exists y_1 \ldots \exists y_n (\bigwedge_{1 \leq j \leq n} y_j \perp_{\vec{z}_j} y_j \wedge \theta)$$

where $\vec{z}_j y_j$, for $1 \leq j \leq n$, is a sequence of at most $k+1$ different variables.   ◀

We will next show the other direction.

▶ **Lemma 28.** *Let $\vec{b} \perp_{\vec{a}} \vec{c}$ be an independence atom where $\vec{a}$, $\vec{b}$ and $\vec{c}$ are tuples of variables. If $\vec{b_0}$ lists the variables in $\vec{b} - \vec{a} \cup \vec{c}$, $\vec{c_0}$ lists the variables in $\vec{c} - \vec{a} \cup \vec{b}$, and $\vec{d}$ lists the variables in $\vec{b} \cap \vec{c} - \vec{a}$, then*

$$\vec{b} \perp_{\vec{a}} \vec{c} \equiv \vec{b_0} \perp_{\vec{a}} \vec{c_0} \wedge \bigwedge_{d \in \vec{d}} =(\vec{a}, d).$$

Now we can prove the following proposition. In the proof we will present a translation from independence logic to ESO, where independence atoms are coded by relation variables preserving the arity of the atoms. Note that the translation presented in [15] does not preserve this property.

▶ **Proposition 29.** $\mathrm{FO}(\perp_c)(k\text{-ind}) \leq \mathrm{ESO}_f(k\text{-ary})$.

Proof. Let $\phi \in \mathrm{FO}(\perp_c)(k\text{-ind})$. By Theorem 15 we may assume that ϕ is in prenex normal form $Q^1 x_1 \ldots Q^n x_n \theta$ where θ is a quantifier-free formula. By Lemma 28 we may assume that each independence atom in θ is either of the form $=(\vec{z}, y)$ or $\vec{b} \perp_{\vec{a}} \vec{c}$ where
- y is not listed in $\vec{z}$,
- $\vec{a}$, $\vec{b}$ and $\vec{c}$ do not share any variables,
- $|\vec{z}| \leq k$ and $|\vec{a}\vec{b}\vec{c}| \leq k+1$.

Let us next consider the subformulas of θ. We will enumerate the subformulas of θ by $\theta_{\vec{i}}$ where $\vec{i}$ is a binary sequence encoding the location of the subformula in θ. Let $\theta_\lambda := \theta$ where λ is the empty sequence. If $\theta_{\vec{i}}$ is a conjunction (or a disjunction), then we denote its conjuncts (or the disjuncts) as $\theta_{\vec{i}0}$ and $\theta_{\vec{i}1}$. Now let $S := \{\vec{i} \mid \theta_{\vec{i}} \text{ is a subformula of } \theta\}$, and let D and I be the subsets of S consisting of sequences $\vec{i}$ for which $\theta_{\vec{i}}$ is a dependence atom or an independence atom, respectively. Let $\leq$ be a partial order in S where $\vec{i} \leq \vec{j}$ if $\vec{i}\vec{k} = \vec{j}$ for some binary $\vec{k}$. Then $\vec{i} \leq \vec{j}$ if and only if $\theta_{\vec{j}}$ is a subformula of $\theta_{\vec{i}}$.

Next we will define a $\Phi \in \mathrm{ESO}_f(k\text{-ary})$ equivalent to ϕ. First we define $\varphi_{\vec{i}}$ for each $\vec{i} \in S$ inductively as follows:
- $\varphi_{\vec{i}} := \theta_{\vec{i}}$ if $\theta_{\vec{i}}$ is a first-order atom,

- $\varphi_{\vec{i}} := S_{\vec{i}}(\vec{a}\vec{b}) \wedge T_{\vec{i}}(\vec{a}\vec{c})$ if $\theta_{\vec{i}}$ is $\vec{b} \perp_{\vec{a}} \vec{c}$,
- $\varphi_{\vec{i}} := f_{\vec{i}}(\vec{z}) = y$ if $\theta_{\vec{i}}$ is $=(\vec{z}, y)$,
- $\varphi_{\vec{i}} := \varphi_{\vec{i}0} \wedge \varphi_{\vec{i}1}$ if $\theta_{\vec{i}}$ is $\theta_{\vec{i}0} \wedge \theta_{\vec{i}1}$,
- $\varphi_{\vec{i}} := \varphi_{\vec{i}0} \vee \varphi_{\vec{i}1}$ if $\theta_{\vec{i}}$ is $\theta_{\vec{i}0} \vee \theta_{\vec{i}1}$.

Now let $\varphi := \varphi_{\lambda}$. Then φ is a quantifier-free first-order formula sharing the structure of θ where the dependence and independence atoms are interpreted using new function symbols $f_{\vec{i}}$ and relation symbols $S_{\vec{i}}$ and $T_{\vec{i}}$, respectively. Let $\vec{z}_{\vec{i}}$, for $\vec{i} \in I$, list the variables in $\{x_1, \ldots, x_n\} \setminus \mathrm{Fr}(\theta_{\vec{i}})$. In the following, for example, $\exists (S_{\vec{i}})_{\vec{i} \in I}$ denotes the prefix $\exists S_{\vec{i}_1} \ldots \exists S_{\vec{i}_m}$ where $\vec{i}_1, \ldots, \vec{i}_m$ enumerates I. So let us define Φ as

$$\exists (S_{\vec{i}})_{\vec{i} \in I} (T_{\vec{i}})_{\vec{i} \in I} (f_{\vec{i}})_{\vec{i} \in D} (Q^1 x_1 \ldots Q^n x_n \varphi \wedge \Omega) \tag{30}$$

where

$$\Omega := \bigwedge_{\vec{i} \in I} [\forall \vec{a}\vec{b}\vec{c}(S_{\vec{i}}(\vec{a}\vec{b}) \wedge T_{\vec{i}}(\vec{a}\vec{c})) \to \exists \vec{z}_{\vec{i}}(\bigwedge_{\vec{j} \leq \vec{i}} \varphi_{\vec{j}} \wedge Q^1 x_1' \ldots Q^n x_n'(\varphi' \wedge \chi))] \tag{31}$$

where $\varphi' := \varphi(x_1'/x_1) \ldots (x_n'/x_n)$ and

$$\chi := \bigwedge_{\substack{1 \leq k \leq n \\ Q^k = \exists}} (x_1 = x_1' \wedge \ldots \wedge x_{k-1} = x_{k-1}') \to x_k = x_k'. \tag{32}$$

The idea behind Φ is that the relation variables $S_{\vec{i}}$ and $T_{\vec{i}}$, for $\vec{i} \in I$, encode a subteam $X_{\vec{i}}$ that satisfies $\vec{b} \perp_{\vec{a}} \vec{c}$. Then Ω will ensure that for each $s, s' \in X_{\vec{i}}$ with $s(\vec{a}) = s'(\vec{a})$ there is s'' corresponding to the values of $\vec{a}\vec{b}\vec{c}$ and $\vec{z}_{\vec{i}}$ such that $s''(\vec{a}\vec{b}\vec{c}) = s(\vec{a}\vec{b})s'(\vec{b})$. The variables $x_{\vec{i}}'$ and χ will ensure that $s'' \in X_{\vec{i}}$. We will now prove that

$$\mathcal{M} \models \phi \Leftrightarrow \mathcal{M} \models \Phi.$$

Only if-part: Assume that $\mathcal{M} \models \phi$. Then there are functions

$$F_i : X[F_1/x_1] \ldots [F_{i-1}/x_{i-1}] \to \mathcal{P}(M) \setminus \{\emptyset\},$$

for $1 \leq i \leq n$, such that

$$\mathcal{M} \models_Y \theta$$

when $Y := \{\emptyset\}[F_1/x_1] \ldots [F_n/x_n]$. Note that $F_i(s) = M$ if $Q^i = \forall$.

Let us then construct teams $Y_{\vec{i}}$, for $\vec{i} \in S$, such that $\mathcal{M} \models_{Y_{\vec{i}}} \theta_{\vec{i}}$, as follows. Let $Y_{\lambda} := Y$.

- Assume that $\mathcal{M} \models_{Y_{\vec{i}}} \theta_{\vec{i}}$ where $\theta_{\vec{i}} = \theta_{\vec{i}0} \wedge \theta_{\vec{i}1}$. Then $Y_{\vec{i}0} := Y_{\vec{i}}$ and $Y_{\vec{i}1} := Y_{\vec{i}}$.
- Assume that $\mathcal{M} \models_{Y_{\vec{i}}} \theta_{\vec{i}}$ where $\theta_{\vec{i}} = \theta_{\vec{i}0} \vee \theta_{\vec{i}1}$. Then choose $Y_{\vec{i}0} \cup Y_{\vec{i}1} = Y_{\vec{i}}$ so that $\mathcal{M} \models_{Y_{\vec{i}0}} \theta_{\vec{i}0}$ and $\mathcal{M} \models_{Y_{\vec{i}1}} \theta_{\vec{i}1}$.

We then note that

$$\mathcal{M} \models_{Y_{\vec{i}}} \vec{b} \perp_{\vec{a}} \vec{c} \quad \text{if} \quad \theta_{\vec{i}} \text{ is } \vec{b} \perp_{\vec{a}} \vec{c}, \tag{33}$$

$$\mathcal{M} \models_{Y_{\vec{i}}} =(\vec{z}, y) \quad \text{if} \quad \theta_{\vec{i}} \text{ is } =(\vec{z}, y). \tag{34}$$

Now, for $\theta_{\vec{i}}$ of the form $\vec{b} \perp_{\vec{a}} \vec{c}$, the interpretations of $S_{\vec{i}}$ and $T_{\vec{i}}$ will be the following:

$$S_{\vec{i}}^{\mathcal{M}} := \{s(\vec{a}\vec{b}) \mid s \in Y_{\vec{i}}\},$$

$$T_{\vec{i}}^{\mathcal{M}} := \{s(\vec{a}\vec{c}) \mid s \in Y_{\vec{i}}\}.$$

For $\theta_{\vec{i}}$ of the form $=(\vec{z}, y)$ we interpret $f_{\vec{i}}$ as follows:

$$f_{\vec{i}}^{\mathcal{M}}(\vec{a}) := \begin{cases} b & \text{if } s(\vec{z}y) = \vec{a}b \text{ for some } s \in Y_{\vec{i}}, \\ 0 & \text{otherwise} \end{cases}$$

where $0 \in M$ is arbitrary. Now $f_{\vec{i}}$ is well defined by (34). Let then $\mathcal{M}^* := (\mathcal{M}, \vec{S}^{\mathcal{M}}, \vec{T}^{\mathcal{M}}, \vec{f}^{\mathcal{M}})$. We will show that

$$\mathcal{M}^* \models Q^1 x_1 \dots Q^n x_n \varphi \wedge \Omega.$$

Consider the first conjunct. For each x_i with $Q^i = \exists$ we can choose a value for it so that the values of $x_1, \dots, x_i$ agree with some $s \in Y$. Thus it suffices to show that, for $s \in Y$,

$$\mathcal{M}^* \models_s \varphi.$$

Since φ is a first-order formula, by Theorem 4 it suffices to show that $\mathcal{M}^* \models_Y \varphi$. This can be done inductively: For each atomic $\varphi_{\vec{i}}$, $\mathcal{M}^* \models_{Y_{\vec{i}}} \varphi_{\vec{i}}$ by the definitions. If $\mathcal{M}^* \models_{Y_{\vec{i}0}} \varphi_{\vec{i}0}$ and $\mathcal{M}^* \models_{Y_{\vec{i}1}} \varphi_{\vec{i}1}$, and $\varphi_{\vec{i}}$ is either disjunction or conjunction of $\varphi_{\vec{i}0}$ and $\varphi_{\vec{i}1}$, then $\mathcal{M}^* \models_{Y_{\vec{i}}} \varphi_{\vec{i}}$ by the construction of $Y_{\vec{i}}$. This concludes the claim and thus the first conjunct part.

Next we will to show that $\mathcal{M}^* \models \Omega$ where Ω is the formula

$$\bigwedge_{\vec{i} \in I} [\forall \vec{a}\vec{b}\vec{c}(S_{\vec{i}}(\vec{a}\vec{b}) \wedge T_{\vec{i}}(\vec{a}\vec{c})) \rightarrow \exists \vec{z}_{\vec{i}} (\bigwedge_{\vec{j} \leq \vec{i}} \varphi_{\vec{j}} \wedge Q^1 x_1' \dots Q^n x_n' (\varphi' \wedge \chi))].$$

Let $\vec{i} \in I$ and assume that $\theta_{\vec{i}} = \vec{b} \perp_{\vec{a}} \vec{c}$. Let $\vec{\alpha}\vec{\beta}\vec{\gamma}$ be such that $\vec{\alpha}\vec{\beta} \in S_{\vec{i}}^{\mathcal{M}}$ and $\vec{\alpha}\vec{\gamma} \in T_{\vec{i}}^{\mathcal{M}}$. Then there are $s, s' \in Y_{\vec{i}}$ such that $s(\vec{a}\vec{b}) = \vec{\alpha}\vec{\beta}$ and $s'(\vec{a}\vec{c}) = \vec{\alpha}\vec{\gamma}$. By (33) we can choose $s'' \in Y_{\vec{i}}$ such that $s''(\vec{a}\vec{b}\vec{c}) = s(\vec{a}\vec{b})s'(\vec{c})$. Let us then choose the values for $\vec{z}_{\vec{i}}$ according to s''. Then all the values of $x_1, \dots, x_n$ agree with s''. Now, since $\mathcal{M}^* \models_{Y_{\vec{j}}} \varphi_{\vec{j}}$ for all $\vec{j}$, and $s'' \in Y_{\vec{j}}$ for $\vec{j} \leq \vec{i}$, it follows by Theorem 4

$$\mathcal{M}^* \models_{s''} \bigwedge_{\vec{j} \leq \vec{i}} \varphi_{\vec{j}}.$$

Now it suffices to show that

$$\mathcal{M}^* \models_{s''} Q^1 x_1' \dots Q^n x_n' (\varphi' \wedge \chi).$$

For each x_i' with $Q^i = \exists$ we choose a value for it so that, for some $t \in Y$, the values of $x_1', \dots, x_i'$ are $t(x_1), \dots, t(x_i)$. In particular, if the values of $x_1', \dots, x_{i-1}'$ agree with s'', then we choose x_i' according to s'' also. Let s^* be an extension of s'' which is constructed according to these rules. Now using the fact that $\mathcal{M}^* \models_t \varphi$ for all $t \in Y$, and the way s^* was chosen, we get

$$\mathcal{M}^* \models_{s^*} \varphi' \wedge \chi.$$

Hence $\mathcal{M} \models \Phi$. This concludes the *only if-part*.

If-part: Assume that $\mathcal{M} \models \Phi$. Then we can find interpretations $\vec{S}^{\mathcal{M}}$, $\vec{T}^{\mathcal{M}}$ and $\vec{f}^{\mathcal{M}}$ such that

$$\mathcal{M}^* \models Q^1 x_1 \dots Q^n x_n \varphi \wedge \Omega$$

when $\mathcal{M}^* := (\mathcal{M}, \vec{S}^{\mathcal{M}}, \vec{T}^{\mathcal{M}}, \vec{f}^{\mathcal{M}})$. Consider the usual semantic game for first-order logic where player $\exists$ plays the role of verifier and player $\forall$ plays the role of falsifier. Then there is a winning strategy for player $\exists$ in the semantic game for $Q^1 x_1 \dots Q^n x_n \varphi \wedge \Omega$ over $\mathcal{M}^*$. Let Y consist of assignments $s : \{x_1, \dots, x_n\} \rightarrow M$ corresponding to every possible play of $x_1, \dots, x_n$ where player $\exists$ follows her winning strategy. Analogously, let Y' consist of

assignments $s : \{x_1, \ldots, x_n\} \to M$ that correspond to every possible play of $x'_1, \ldots, x'_n$ where player $\exists$ follows her winning strategy. Let $X := Y \cup Y'$. We will show that

$$\mathcal{M} \models_X \theta.$$

We know that $\mathcal{M}^* \models_X \varphi$. Let us now define $X_{\vec{i}}$, for $\vec{i} \in S$, as follows. Recall that $\varphi_\lambda = \varphi$ where λ is the empty sequence. We also let $X_\lambda := X$.

- If $\mathcal{M}^* \models_{X_{\vec{i}}} \varphi_{\vec{i}}$ where $\varphi_{\vec{i}} = \varphi_{\vec{i}0} \wedge \varphi_{\vec{i}1}$, then we let $X_{\vec{i}0} := X_{\vec{i}}$ and $X_{\vec{i}1} := X_{\vec{i}}$.
- If $\mathcal{M}^* \models_{X_{\vec{i}}} \varphi_{\vec{i}}$ where $\varphi_{\vec{i}} = \varphi_{\vec{i}0} \vee \varphi_{\vec{i}1}$, then we let $X_{\vec{i}0} := \{s \in X_{\vec{i}} \mid \mathcal{M}^* \models_s \varphi_{\vec{i}0}\}$ and $X_{\vec{i}1} := \{s \in X_{\vec{i}} \mid \mathcal{M}^* \models_s \varphi_{\vec{i}1}\}$.

From the construction it follows that $\mathcal{M}^* \models_{X_{\vec{i}}} \varphi_{\vec{i}}$, for $\vec{i} \in S$, and that $X_{\vec{i}0} \cup X_{\vec{i}1} = X_{\vec{i}}$ if $\varphi_{\vec{i}}$ is a disjunction. We will now show that for each atomic $\theta_{\vec{i}}$, $\mathcal{M} \models_{X_{\vec{i}}} \theta_{\vec{i}}$:

1. If $\theta_{\vec{i}}$ is a first-order atom, then the claim follows from $\theta_{\vec{i}} = \varphi_{\vec{i}}$.
2. If $\theta_{\vec{i}}$ is $=(\vec{z}, y)$, then the claim follows from $\mathcal{M}^* \models_{X_{\vec{i}}} f_{\vec{i}}(\vec{z}) = y$.
3. Assume that $\theta_{\vec{i}}$ is $\vec{b} \perp_{\vec{a}} \vec{c}$. Then $\mathcal{M}^* \models_{X_{\vec{i}}} S_{\vec{i}}(\vec{a}\vec{b}) \wedge T_{\vec{i}}(\vec{a}\vec{c})$. Let $s, s' \in X_{\vec{i}}$ be such that $s(\vec{a}) = s'(\vec{a})$. We have to show that there is $s'' \in X_{\vec{i}}$ such that $s''(\vec{a}\vec{b}\vec{c}) = s(\vec{a})s(\vec{b})s'(\vec{c})$. Now $\mathcal{M}^* \models \Omega$, so consider a play in the semantic game where player $\forall$ chooses first the conjunct with index $\vec{i}$ from Ω, and then chooses $s(\vec{a})s(\vec{b})s'(\vec{c})$ as values for $\vec{a}\vec{b}\vec{c}$. Since $s(\vec{a})s(\vec{b}) \in S_{\vec{i}}^{\mathcal{M}}$ and $s(\vec{a})s'(\vec{c}) \in T_{\vec{i}}^{\mathcal{M}}$, then player $\exists$ plays according to her strategy and chooses values for $\vec{z}_{\vec{i}}$ so that

$$\mathcal{M}^* \models_{s''} \bigwedge_{\vec{j} \leq \vec{i}} \varphi_{\vec{j}} \wedge Q^1 x'_1 \ldots Q^n x'_n (\varphi' \wedge \chi)$$

where s'' is the assignment agreeing with the chosen values for $\vec{a}\vec{b}\vec{c}$ and $\vec{z}_{\vec{i}}$. Now we let player $\forall$ play each x'_i with $Q^i = \forall$ as $s''(x_i)$. Then because of χ (defined in (32)) player $\exists$ must also play each x'_i with $Q^i = \exists$ as $s''(x_i)$. Hence s'' corresponds to a play of $x'_1, \ldots, x'_n$, and thus $s'' \in X$.

Since $\mathcal{M}^* \models_{s''} \bigwedge_{\vec{j} \leq \vec{i}} \varphi_{\vec{j}}$, it is a straightforward induction to show that $s'' \in X_{\vec{i}}$. This concludes the step 3.

Now using the previous, a straightforward backward induction shows that $\mathcal{M} \models_X \theta$. It then suffices to show that there are functions

$$F_i : \{\emptyset\}[F_1/x_1] \ldots [F_{i-1}/x_{i-1}] \to \mathcal{P}(M) \setminus \{\emptyset\}$$

such that $F_i(s) = M$ if $Q^i = \forall$, and that

$$X = \{\emptyset\}[F_1/x_1] \ldots [F_n/x_n].$$

We define these functions inductively so that $\{\emptyset\}[F_1/x_1] \ldots [F_i/x_i] = X \restriction \{x_1, \ldots, x_i\}$, for $1 \leq i \leq n$. Assume that we have defined $F_1, \ldots, F_i$ successfully. We will define F_{i+1} as wanted. Assume first that $Q^{i+1} = \exists$. Then for $s \in \{\emptyset\}[F_1/x_1] \ldots [F_i/x_i]$, we let

$$F_{i+1}(s) = \{t(x_{i+1}) \mid t \in X, t \restriction \{x_1, \ldots, x_i\} = s\}.$$

By the induction assumption $F_{i+1}(s)$ is non-empty, though it may not be singleton in case there are multiple plays where values of $x_1, \ldots, x_i$ (or $x'_1, \ldots, x'_i$) agree with s. We note that

$$\{\emptyset\}[F_1/x_1] \ldots [F_{i+1}/x_{i+1}] = X \restriction \{x_1, \ldots, x_{i+1}\}.$$

Assume then that $Q^{i+1} = \forall$. For $s \in \{\emptyset\}[F_1/x_1] \ldots [F_i/x_i]$, we let $F_{i+1}(s) = \mathcal{P}(M)$ and note that

$$X \restriction \{x_1, \ldots, x_{i+1}\} \subseteq \{\emptyset\}[F_1/x_1] \ldots [F_{i+1}/x_{i+1}].$$

For the other direction, assume that $s \in \{\emptyset\}[F_1/x_1]\dots[F_i/x_i]$ and let $a \in M$. We show that $s(a/x) \in X \upharpoonright \{x_1,\dots,x_{i+1}\}$. By the induction assumption $s \in X \upharpoonright \{x_1,\dots,x_i\}$, and thus there is a play of $x_1,\dots,x_n$ (or $x_1',\dots,x_n'$) that agrees with s in the first i variables. Let s' be the assignment corresponding to this play. Now instead of choosing $s'(x_{i+1})$ (or $s'(x_{i+1}')$) at move $i+1$, player $\forall$ can choose a for x_{i+1} (or for x_{i+1}'). Let t be an assignment that corresponds to some play with these moves for the first $i+1$ variables. Then $t \in X$ and $t \upharpoonright \{x_1,\dots,x_{i+1}\} = s(a/x_{i+1})$. This concludes the proof, and thus also the *only if-part*.

Note that in Φ each function or relation variable has an arity at most k. This concludes the proof. ◀

▶ **Theorem 35.** $\mathrm{ESO}_f(k\text{-ary}) = \mathrm{FO}(\perp_c)(k\text{-ind})$.

Proof. Follows from Propositions 27 and 29. ◀

This gives us immediately a corollary regarding inclusion logic. Recall that $\mathrm{FO}(\subseteq)(k\text{-inc})$ denotes the class of inclusion logic sentences in which inclusion atoms of width at most k (i.e. atoms of the form $\vec{a} \subseteq \vec{b}$ where $|\vec{a}| = |\vec{b}| \le k$) may appear.

▶ **Theorem 36.** *Assume* $k \ge 2$. *Then* $\mathrm{FO}(\subseteq)(k\text{-inc}) \le \mathrm{ESO}_f(k\text{-ary})$.

Proof. Using item 1 of Theorem 12, we first translate inclusion logic sentences to independence logic, and then apply Proposition 29. It is easy to check that this translation takes us to $\mathrm{ESO}_f(k'\text{-ary})$, where $k' = \max\{k, 2\}$. ◀

There is no hope of proving the other direction, since, e.g., even cardinality cannot be expressed in $\mathrm{FO}(\subseteq)$ [3], but it is expressible in $\mathrm{ESO}_f(1\text{-ary})$. Next we will show that $\mathrm{ESO}_f(k\text{-ary}) \le \mathrm{FO}(\perp)(2k+2\text{-ind})$.

▶ **Theorem 37.** $\mathrm{ESO}_f(k\text{-ary}) \le \mathrm{FO}(\perp)(2k+1\text{-ind}) \le \mathrm{ESO}_f(2k+1\text{-ary})$.

Proof. For the first inequality, note that $\mathrm{ESO}_f(k\text{-ary}) = \mathcal{D}(k\text{-dep})$ by [7], and $\mathcal{D}(k\text{-dep}) \le \mathrm{FO}(\perp)(2k+1\text{-ind})$ by Theorem 13. The second inequality follows from Theorem 35. ◀

4.2   ∀-hierarchies

In this section, we will examine the fragments $\mathrm{FO}(\mathcal{C})(k\forall)$. We will prove that, contrary to the case of the fragments $\mathcal{D}(k\forall)$ [7], the following holds:
1. If $\{\perp, \subseteq\} \subseteq \mathcal{C}$ then the hierarchy collapses at level 1: $\mathrm{FO}(\mathcal{C}) = \mathrm{FO}(\mathcal{C})(1\forall)$;
2. If $\perp \in \mathcal{C}$ then it collapses at level 2: $\mathrm{FO}(\mathcal{C}) = \mathrm{FO}(\mathcal{C})(2\forall)$.

We will use the following result from [21]:

▶ **Proposition 38.** *Let* ϕ *be a* $\mathrm{FO}(\perp)$ *sentence. Then* ϕ *is equivalent to an formula of the form* $\forall\vec{x}\exists\vec{y}(\theta \wedge \chi)$, *where* θ *is a conjunction of pure independence atoms and* χ *is first-order and quantifier-free.*

Since, as we saw in the Preliminaries, we can define inclusion atoms and conditional independence atoms in terms of pure independence atoms, it follows at once that any sentence of $\mathrm{FO}(=(\dots), \perp_c, \subseteq)$ is equivalent to some sentence of the above form.

Using this, we will prove that

▶ **Theorem 39.** $\mathrm{FO}(\perp, \subseteq)(1\forall) = \mathrm{FO}(=(\dots), \perp_c, \subseteq)$.

Proof. Let $\phi \in \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$. We will show that there exists a $\phi' \in \mathrm{FO}(\perp, \subseteq)(1\forall)$ equivalent to it. As we said, we can assume that ϕ is of the form $\forall x_1 \ldots \forall x_m \exists x_{m+1} \ldots \exists x_{m+n}(\theta \wedge \chi)$, where θ is a conjunction of pure independence atoms and χ is first-order and quantifier-free. Let us then define ϕ' as

$$\forall x_1 \exists x_2 \ldots \exists x_m \exists x_{m+1} \ldots \exists x_{m+n}(\bigwedge_{2 \leq i \leq m}(x_1 \subseteq x_i \wedge x_1 \ldots x_{i-1} \perp x_i) \wedge \theta \wedge \chi).$$

We claim that ϕ' is equivalent to ϕ. Assume first that $\mathcal{M} \models \phi$. Then there are, for $m+1 \leq i \leq m+n$, functions

$$F_i : \{\emptyset\}[M/x_1] \ldots [M/x_n][F_{m+1}/x_{m+1}] \ldots [F_{i-1}/x_{i-1}] \to \mathcal{P}(M) \setminus \{\emptyset\}$$

such that $\mathcal{M} \models_X \theta \wedge \chi$ when $X := [M/x_1] \ldots [M/x_n][F_{m+1}/x_{m+1}] \ldots [F_{m+n}/x_{m+n}]$. Let F_i, for $2 \leq i \leq m$, be the constant function mapping each assignment to M. Then

$$X = \{\emptyset\}[M/x_1][F_2/x_2] \ldots [F_{m+n}/x_{m+n}].$$

Clearly $\mathcal{M} \models_X \bigwedge_{2 \leq i \leq m}(x_1 \subseteq x_i \wedge x_1 \ldots x_{i-1} \perp x_i)$, and hence $\mathcal{M} \models \phi'$.

For the other direction, assume that $\mathcal{M} \models \phi'$. Then there are, for $2 \leq i \leq m+n$, functions

$$F_i : \{\emptyset\}[M/x_1][F_2/x_2] \ldots [F_{i-1}/x_{i-1}] \to \mathcal{P}(M) \setminus \{\emptyset\}$$

such that $\mathcal{M} \models_X \bigwedge_{2 \leq i \leq m}(x_1 \subseteq x_i \wedge x_1 \ldots x_{i-1} \perp x_i)$ when

$$X := \{\emptyset\}[M/x_1][F_2/x_2] \ldots [F_{m+n}/x_{m+n}].$$

Define, for $2 \leq i \leq m$,

$$X_i := \{\emptyset\}[M/x_1][F_2/x_2] \ldots [F_i/x_i]$$

and

$$Y_i := \{\emptyset\}[M/x_1][M/x_2] \ldots [M/x_i].$$

It suffices to show that $X_i = Y_i$ for $2 \leq i \leq m$.

First let us prove the claim for $i = 2$. Let $s \in Y_2$. It suffices to show that $s \in X_2$. By Proposition 7, $\mathcal{M} \models_{X_2} x_1 \subseteq x_2 \wedge x_1 \perp x_2$. Let $s' \in X_2$ be such that $s'(x_1) = s(x_2)$. Since $\mathcal{M} \models_{X_2} x_1 \subseteq x_2$, we can find a $t \in X_2$ such that $t(x_2) = s'(x_1)$. Now let $t' \in X_2$ be such that $t'(x_1) = s(x_1)$. Because $\mathcal{M} \models_{X_2} x_1 \perp x_2$, we can find a $t'' \in X_2$ such that $t''(x_1) = t'(x_1)$ and $t''(x_2) = t(x_2)$. Then $t'' = s$ which concludes the claim for $i = 2$.

The induction step is proved analogously. This concludes the claim and the proof. ◄

Let us now prove our second claim.

▶ **Theorem 40.** $\mathrm{FO}(\perp)(2\forall) = \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$.

Proof. Let $\phi \in \mathrm{FO}(=(\ldots), \perp_c, \subseteq)$. Again, we can assume that ϕ is of the form $\forall \vec{x} \psi$, where $\vec{x} = x_1 \ldots x_n$ and ψ is of the form $\exists \vec{y} \theta$ for θ quantifier-free and in $\mathrm{FO}(\perp)$. Let now p, q be two variables not occurring in ϕ. We state that ϕ is equivalent to

$$\phi^* = \forall p \forall q \exists \vec{x} \left(\left(p = q \to \bigwedge_{i=1}^{n} x_i = p \right) \wedge \bigwedge_{i=1}^{n-1}(x_1 \ldots x_i \perp x_{i+1}) \wedge \psi \right).$$

Indeed, let $\mathcal{M}$ be a model and $X = \{\emptyset\}[M/p][M/q]$, and let the tuple of (nondeterministic) choice functions $\vec{U}$ for $\vec{x}$ be such that

$$\vec{U}(s) = \begin{cases} (m, \ldots, m) & \text{if } s(p) = s(q) = m; \\ M^n & \text{otherwise} \end{cases}$$

and let $Y = X[\vec{U}/\vec{x}]$. It is obvious that $\mathcal{M} \models_Y (p = q \rightarrow \bigwedge_{i=1}^n x_i = p)$; and $\mathcal{M} \models_Y \psi$, because $Y(\vec{x}) = M^n$ and p, q do not occur in ψ. Finally, it is also true that Y satisfies all independence atoms $x_1 \ldots x_i \perp x_{i+1}$, since $Y(x_1 \ldots x_i x_{i+1}) = M^{i+1}$ (assuming that our model contains two distinct elements). Therefore $\mathcal{M} \models \phi^*$, as required.

Conversely, suppose $\mathcal{M} \models \phi^*$ when there exists a $\vec{U}$ such that, for $Y = \{\emptyset\}[M/pq][\vec{U}/\vec{x}]$, $\mathcal{M} \models_Y (p = q \rightarrow \bigwedge_{i=1}^n x_i = p) \wedge \bigwedge_{i=1}^{n-1}(x_1 \ldots x_i \perp x_{i+1}) \wedge \psi$. We will show that $Y(x_1 \ldots x_n)$ is M^n, that is, that all possible tuples $m_1 \ldots m_n$ of elements of our models are possible values for $x_1 \ldots x_n$ in Y.

First of all, let us observe that for all $m \in M$ there exists a $h^m \in Y$ such that $h^m(x_i) = m$ for all i. Indeed, we can find a $s^m \in X$ such that $s^m(p) = s^m(q) = m$ and then pick an arbitrary $h^m \in s^m[\vec{U}/\vec{x}] \subseteq Y$. Since $\mathcal{M} \models_Y p = q \rightarrow \bigwedge_i x_i = p$, we have at once that $h^m(x_i) = h^m(p) = m$, as required.

Now we prove, by induction on $i = 1 \ldots n$, that there exists a $h_i \in Y$ such that $h_i(x_1 \ldots x_i) = m_1 \ldots m_i$.

Base Case: Let h_1 be $h^{m_1} \in Y$. Then $h^{m_1}(x_1) = m_1$, as required.

Induction Case: Suppose that $h_i(x_1 \ldots x_i) = m_1 \ldots m_i$, and consider $h^{m_{i+1}}$. As we saw, $h^{m_{i+1}} \in Y$ and $h^{m_{i+1}}(x_{i+1}) = m_{i+1}$. But $\mathcal{M} \models_Y x_1 \ldots x_i \perp x_{i+1}$; and therefore there exists a $h_{i+1} \in Y$ with $h_{i+1}(x_1 \ldots x_i) = h_i(x_1 \ldots x_i) = m_1 \ldots m_i$ and $h_{i+1}(x_{i+1}) = h^{m_{i+1}}(x_{i+1}) = m_{i+1}$. Hence, $h_{i+1}(x_1 \ldots x_{i+1}) = m_1 \ldots m_{i+1}$.

In particular, this implies that $h_n(x_1 \ldots x_n) = m_1 \ldots m_n$; and since we started from an arbitrary choice of $m_1 \ldots m_n$, we can conclude that $Y(\vec{x}) = M^{|\vec{x}|}$. But then the restriction of Y to $\vec{x}$ is precisely $\{\emptyset\}[M/\vec{x}]$; and since $\mathcal{M} \models_Y \psi$, by locality we have that $\mathcal{M} \models \forall \vec{x} \psi$, as required. ◀

Conclusion

In this paper, we examined the expressive power of fragments of inclusion and independence logic obtained by restricting the arity of non first-order atoms or the number of universal quantifiers. For the first kind of restriction, we adapted and extended the hierarchy theorems of [7] to this new setting; but for the second kind of restriction, we showed that the hierarchy collapses at a very low level if our logic contains at least pure independence atoms.

A question which is still open is whether the fragments $\mathrm{FO}(\subseteq)(k\forall)$ of inclusion logic give rise to an infinite expressivity hierarchy. Another issue that requires further investigation is to which degree our results can be adapted to the case of strict semantics. The exact nature of the relationship between strict and lax semantics is a matter which is of no small interest for the further development of the area, and a comparison of the properties of our fragments in these two settings might prove itself of great value.

───── **References** ─────

1   Samson Abramsky and Jouko Väänänen. From IF to BI. *Synthese*, 167:207–230, 2009. 10.1007/s11229-008-9415-6.

2   Miklos Ajtai. Σ_1^1-formulae on finite structures. *Ann. Pure Appl. Logic*, 4(1):1 – 48, 1983.

3   Pietro Galliani Lauri Hella. Inclusion logic and fixpoints. Manuscript, 2013.

4   William W. Armstrong. Dependency Structures of Data Base Relationships. In *Proc. of IFIP World Computer Congress*, pages 580–583, 1974.

5   Marco A. Casanova, Ronald Fagin, and Christos H. Papadimitriou. Inclusion dependencies and their interaction with functional dependencies. In *Proceedings of the 1st ACM SIGACT-*

SIGMOD symposium on Principles of database systems, PODS '82, pages 171–176, New York, NY, USA, 1982. ACM.

6 Stephen A. Cook. A hierarchy for nondeterministic time complexity. In *Conference Record, Fourth Annual ACM Symposium on Theory of Computing*, pages 187–192. ACM, 1972.

7 Arnaud Durand and Juha Kontinen. Hierarchies in dependence logic. *ACM Transactions on Computational Logic (TOCL)*, 13(4):31, 2012.

8 Herbert B. Enderton. Finite partially-ordered quantifiers. *Mathematical Logic Quarterly*, 16(8):393–397, 1970.

9 Ronald Fagin. A normal form for relational databases that is based on domains and keys. *ACM Transactions on Database Systems*, 6:387–415, September 1981.

10 Pietro Galliani. *The Dynamics of Imperfect Information*. PhD thesis, University of Amsterdam, September 2012.

11 Pietro 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.

12 Dan Geiger, Azaria Paz, and Judea Pearl. Axioms and algorithms for inferences involving probabilistic independence. *Information and Computation*, 91(1):128–141, 1991.

13 Erich Grädel. Model-checking games for logics of imperfect information. *Theoretical Computer Science (to appear)*, 2012.

14 Etienne Grandjean and Frédéric Olive. Graph properties checkable in linear time in the number of vertices. *J. Comput. Syst. Sci.*, 68(3):546–597, 2004.

15 Erich Grädel and Jouko Väänänen. Dependence and independence. *Studia Logica*, 101(2):399–410, 2013.

16 Miika Hannula. Axiomatizing first-order consequences in independence logic. Manuscript, 2012.

17 Leon Henkin. Some Remarks on Infinitely Long Formulas. In *Infinitistic Methods. Proc. Symposium on Foundations of Mathematics*, pages 167–183. Pergamon Press, 1961.

18 Jarmo Kontinen. Coherence and computational complexity of quantifier-free dependence logic formulas. *Studia Logica*, 101(2):267–291, 2013.

19 Pavel Naumov and Brittany Nicholls. R.E. axiomatization of conditional independence. To appear, 2013.

20 Jouko Väänänen. *Dependence Logic*. Cambridge University Press, 2007.

21 Jouko Väänänen and Pietro Galliani. On dependence logic. In Preparation, 2013.

22 Wilbur John Walkoe. Finite partially-ordered quantification. *The Journal of Symbolic Logic*, 35(4):pp. 535–555, 1970.

23 Fan Yang. Expressing Second-order Sentences in Intuitionistic Dependence Logic. *Studia Logica*, 101(2):323–342, 2013.

5 Appendix

This secrtion contains the proof omitted in the main part of the paper.

▶ **Theorem 13.** *For all models $\mathcal{M}$ and teams X*

$$\mathcal{M} \models_X =(\vec{x}, y) \Leftrightarrow \mathcal{M} \models_X \forall \vec{z} \exists w ((\vec{z} = \vec{x} \rightarrow w = y) \wedge \vec{x} y \perp \vec{z} w).$$

Proof. Suppose first that $\mathcal{M} \models_X =(\vec{x}, y)$. Then there exists a function $f : M^{|\vec{x}|} \rightarrow M$ such that $f(s(\vec{x})) = s(y)$ for all $s \in X$. Then for $Y = X[M/\vec{z}]$, define the choice function $H : Y \rightarrow \mathcal{P}(M) \setminus \{\emptyset\}$ so that

$$H(s) = \{f(s(\vec{z}))\}$$

for all $s \in Y$, and let $Z = Y[H/w]$. If we can verify that $\mathcal{M} \models_Z \vec{z} = \vec{x} \to w = y$ and that $\mathcal{M} \models_Z \vec{x}y \perp \vec{z}w$, the left-to-right direction of our proof is done. Now, if $h \in Z$ then $h(w) = f(h(\vec{z}))$ and $h(y) = f(h(\vec{x}))$, and therefore $\mathcal{M} \models_Z \vec{z} = \vec{x} \to w = y$. Furthermore, for $h, h' \in Z$, we have that $h'' = h[h'(\vec{z})/\vec{z}][h'(w)/w] \in Z$, since our choice of w depends only on $\vec{z}$, and therefore $\mathcal{M} \models_Z \vec{x}y \perp \vec{z}w$.

Conversely, suppose that there exists a function $H : X[M/\vec{z}] \to \mathcal{P}(M) \backslash \{\emptyset\}$ such that, for $Z = X[M/\vec{z}][H/w]$, $\mathcal{M} \models_Z \vec{z} = \vec{x} \to w = y \wedge \vec{x}y \perp \vec{z}w$. Now let $s, s' \in X$ be such that $s(\vec{x}) = s'(\vec{x}) = \vec{m}$, let $a = s(y)$ and let $b = s'(y)$: we need to prove that $a = b$.

Take $h \in s[\vec{m}/\vec{z}][H/w] \subseteq Z$: since $\mathcal{M} \models_Z \vec{z} = \vec{x} \to w = y$, we must have that $h(w) = s(y) = a$. Similarly, for $h' \in s'[\vec{m}/\vec{z}][H/w] \subseteq Z$, we must have that $h'(w) = s'(y) = b$. But $\mathcal{M} \models_Z \vec{x}y \perp \vec{z}w$, so there exists a $h'' \in Z$ such that $h''(\vec{x}y) = h(\vec{x}y) = \vec{m}a$ and $h''(\vec{z}w) = h'(\vec{z}w) = \vec{m}b$. Since, again, $\mathcal{M} \models_Z \vec{z} = \vec{x} \to w = y$, the only possibility is that $a = b$, as required. ◀

▶ **Lemma 14.** *Let $\phi, \psi \in \mathrm{FO}(=(\ldots), \perp_{\mathrm{c}}, \subseteq)$ and let x be a variable not occurring free in ψ. Then the following equivalences hold:*

1. $\exists x \phi \wedge \psi \equiv \exists x(\phi \wedge \psi)$,
2. $\exists x \phi \vee \psi \equiv \exists x(\phi \vee \psi)$,
3. $\forall x \phi \wedge \psi \equiv \forall x(\phi \wedge \psi)$,
4. $\forall x \phi \vee \psi \equiv \exists a \exists b \forall x((\phi \wedge a = b) \vee (\psi \wedge a \neq b))$ *where a and b are new variables.*

Proof. The cases 1, 2 and 3 are proved as in Lemma 12 in [16]. We prove number 4. By Proposition 7 it is enough to prove the equivalence for teams X with $\mathrm{Dom}(X) = \mathrm{Fr}(\forall x \phi \vee \psi)$.

Assume first that $\mathcal{M} \models_X \forall x \phi \vee \psi$ and x does not occur free in ψ. Then there are $Y \cup Z = X$ such that $\mathcal{M} \models_{Y[M/x]} \phi$ and $\mathcal{M} \models_Z \psi$. Let $0, 1 \in M$ be distinct. We extend each $s \in X$ with $a \mapsto 0$ and $b \mapsto 0$, for $s \in Y$, and with $a \mapsto 0$ and $b \mapsto 1$, for $s \in Z$, and we let X' consist of these extended assignments. So each $s \in X$ has either one or two extensions in X'. Let $Y' := \{s \in X'[M/x] \mid s(a) = s(b)\}$ and $Z' := \{s \in X'[M/x] \mid s(a) \neq s(b)\}$. Then by Proposition 7, $\mathcal{M} \models_{Y'} \phi \wedge a = b$ and $M \models_{Z'} \psi \wedge a \neq b$. Hence $M \models_{X'[M/x]} (\phi \wedge a = b) \vee (\psi \wedge a \neq b)$, and we conclude that $\mathcal{M} \models_X \exists a \exists b \forall x((\phi \wedge a = b) \vee (\psi \wedge a \neq b))$.

Assume then that $\mathcal{M} \models_X \exists a \exists b \forall x((\phi \wedge a = b) \vee (\psi \wedge a \neq b))$. Let $F_a : X \to \mathcal{P}(M)$ and $F_b : X[F_a/a] \to \mathcal{P}(M)$ be such that if $X' := X[F_a/a][F_b/b][M/x]$, then $\mathcal{M} \models_{X'} (\phi \wedge a = b) \vee (\psi \wedge a \neq b)$. Let $Y' \cup Z' = X'$ be such that $\mathcal{M} \models_{Y'} \phi \wedge a = b$ and $\mathcal{M} \models_{Z'} \psi \wedge a \neq b$. Let $Y := Y' \upharpoonright \mathrm{Dom}(X)$ and $Z := Z' \upharpoonright \mathrm{Dom}(X)$. Then $Y[M/x] = Y' \upharpoonright (\mathrm{Dom}(X) \cup \{x\})$, and thus by Proposition 7 $\mathcal{M} \models_{Y[M/x]} \phi$. Also by Proposition 7 $\mathcal{M} \models_Z \psi$. Since $Y \cup Z = X$, we conclude that $\mathcal{M} \models \forall x \phi \vee \psi$. ◀

▶ **Theorem 25.** *For each natural number n there is a sentence $\phi \in \mathrm{FO}(\subseteq)$ such that for all models $\mathcal{M}$ and teams X where $X \neq \emptyset$ and the variables in $Dom(X)$ do not appear in ϕ,*

$$\mathcal{M} \models_X \phi \text{ if and only if } |X| \geq n.$$

Proof. Let n be a natural number. We may assume that $n \geq 2$ because in the case $n = 1$ we can just choose $\phi := \top$. Let $\vec{x}_i$, for $0 \leq i \leq n-1$, list variables $x_{i,0}, \ldots, x_{i,l}$ where $l = \log(n)$. Let

$$\phi := \exists \vec{x}_0 \ldots \exists \vec{x}_{n-1} (\bigwedge_{0 \leq i \leq n-1} \vec{x}_i \subseteq \vec{x}_0 \wedge \bigwedge_{0 \leq i < j \leq n-1} \vec{x}_i \neq \vec{x}_j)$$

where

$$\vec{x}_i \neq \vec{x}_j := \bigvee_{0 \leq k \leq l} x_{i,k} \neq x_{j,k}.$$

Now ϕ is as wanted:

Assume first that $\mathcal{M} \models_X \phi$. Then there are, for $0 \leq i \leq n-1$, functions

$$F_i : X[F_0/\vec{x}_0]\ldots[F_{i-1}/\vec{x}_{i-1}] \to M^{l+1}$$

such that

$$M \models_{X'} \bigwedge_{0 \leq i \leq n-1} \vec{x}_i \subseteq \vec{x}_0 \wedge \bigwedge_{0 \leq i < j \leq n-1} \vec{x}_i \neq \vec{x}_j \tag{41}$$

when $X' := X[F_0/\vec{x}_0]\ldots[F_{n-1}/\vec{x}_{n-1}]$. Let $s \in X'$ be some arbitrary assignment. From (41) it follows that X' must include assignments s_i, for $0 \leq i \leq n-1$, such that $s_i(\vec{x}_0) = s(\vec{x}_i)$. Also from (41) it follows that $s(\vec{x}_i) \neq s(\vec{x}_j)$, for $0 \leq i < j \leq n-1$. Thus the assignments s_i are distinct and therefore $|X'| \geq n$. Because existential quantification of new variables in strict semantics preserves the cardinality of a team we deduce that $X \geq n$.

Suppose then $X \geq n$. By the assumption $n \geq 2$, and thus we may deduce that $|M| \geq 2$. Let 0 and 1 be two different members of M, and let $\bar{i}$ be the binary representation (of length $l+1$) of i, for $0 \leq i \leq n-1$, in terms of these 0 and 1. Choose then n different assignments $s_0, \ldots, s_{n-1}$ from X. We define, for $0 \leq i \leq n-1$, $F_i : X[F_0/\vec{x}_0]\ldots[F_{i-1}/\vec{x}_{i-1}] \to M^{l+1}$ as follows:

$$F_i(s) := \begin{cases} \overline{j+i} & \text{if } s \upharpoonright \mathrm{Dom}(X) = s_j, \text{ for } 0 \leq j \leq n-1, \\ \bar{i} & \text{otherwise} \end{cases}$$

where $j+i$ is mod n. By the assumption, the variables in $\mathrm{Dom}(X)$ are not listed in $\vec{x}_0 \ldots \vec{x}_{n-1}$, and thus the functions F_i are consistent with the definition of existential quantification for strict semantics. Without the assumption it could be the case that different s_i and s_j would collapse into one assignment in the quantification procedure. Let $X' := X[F_0/\vec{x}_0]\ldots[F_{n-1}/\vec{x}_{n-1}]$. Then s_j, for $0 \leq j \leq n-1$, is extended in X' to

$$s_j(\bar{j}/\vec{x}_0)(\overline{j+1}/\vec{x}_1)\ldots(\overline{j-2}/\vec{x}_{n-2})(\overline{j-1}/\vec{x}_{n-1}),$$

and each $t \in X \setminus \{s_j \mid 0 \leq j \leq n-1\}$ is extended in X' analogously to s_0. So for each $s \in X'$ and $0 \leq i < j \leq n-1$ it holds that $s(\vec{x}_i) \neq s(\vec{x}_j)$. Also

$$\{s(\vec{x}_0) \mid s \in X'\} = \{\bar{i} \mid 0 \leq i \leq n-1\} = \bigcup_{0 \leq i \leq n-1} \{s(\vec{x}_i) \mid s \in X'\},$$

and thus

$$M \models_{X'} \bigwedge_{0 \leq i \leq n-1} \vec{x}_i \subseteq \vec{x}_0 \wedge \bigwedge_{0 \leq i < j \leq n-1} \vec{x}_i \neq \vec{x}_j$$

which concludes the proof.   ◀

▶ **Lemma 28.** *Let $\vec{b} \perp_{\vec{a}} \vec{c}$ be an independence atom where $\vec{a}$, $\vec{b}$ and $\vec{c}$ are tuples of variables. If $\vec{b_0}$ lists the variables in $\vec{b} - \vec{a} \cup \vec{c}$, $\vec{c_0}$ lists the variables in $\vec{c} - \vec{a} \cup \vec{b}$, and $\vec{d}$ lists the variables in $\vec{b} \cap \vec{c} - \vec{a}$, then*

$$\vec{b} \perp_{\vec{a}} \vec{c} \equiv \vec{b_0} \perp_{\vec{a}} \vec{c_0} \wedge \bigwedge_{d \in \vec{d}} =(\vec{a}, d).$$

Proof. Assume that $\mathcal{M} \models_X \vec{b} \perp_{\vec{a}} \vec{c}$. Then clearly $\vec{b_0} \perp_{\vec{a}} \vec{c_0}$. For $\bigwedge_{d \in \vec{d}} =(\vec{a}, d)$, let $d \in \vec{d}$ and $s, s' \in X$ be such that $s(\vec{a}) = s'(\vec{a})$. Then by the assumption there is $s'' \in X$ such that $s''(\vec{a}\vec{b}\vec{c}) = s(\vec{a}\vec{b})s'(\vec{c})$. Because d is listed in both $\vec{b}$ and $\vec{c}$, it follows that $s(d) = s'(d)$.

Suppose then $\mathcal{M} \models_X \vec{b_0} \perp_{\vec{a}} \vec{c_0} \wedge \bigwedge_{d \in \vec{d}} = (\vec{a}, d)$. Let $s, s' \in X$ be such that $s(\vec{a}) = s'(\vec{a})$. By the assumption there is $s'' \in X$ such that $s''(\vec{a}\vec{b_0}\vec{c_0}) = s(\vec{a}\vec{b_0})s'(\vec{c_0})$. We want to show that $s''(\vec{a}\vec{b}\vec{c}) = s(\vec{a}\vec{b})s'(\vec{c})$. Consider first variables x listed in $\vec{b} - \vec{b_0}$. If x is listed in $\vec{a}$, then $s''(x) = s(x)$ as wanted. Assume that x is listed in $\vec{c} - \vec{a}$. Then $x \in \vec{d}$, and thus $s''(x) = s(x)$ follows from $s''(\vec{a}) = s(\vec{a})$.

For variables x is listed in $\vec{c} - \vec{c_0}$ the proof of $s''(x) = s'(x)$ is analogous because $s(\vec{a}) = s'(\vec{a})$. This concludes the proof. ◀