

Hierarchies in inclusion logic with lax semantics*

Miika Hannula**

University of Helsinki

Abstract. We study the expressive power of fragments of inclusion logic under the so-called lax team semantics. The fragments are defined either by restricting the number of universal quantifiers or the arity of inclusion atoms in formulae. In case of universal quantifiers, the corresponding hierarchy collapses at the first level. Arity hierarchy is shown to be strict by relating the question to the study of arity hierarchies in fixed point logics.

1 Introduction

In this article we study the expressive power of inclusion logic ($\text{FO}(\subseteq)$) [4] in the lax team semantics setting. Inclusion logic is a variant of dependence logic ($\text{FO}(=\dots)$) [16] which extends first-order logic with dependence atoms

$$=(x_1, \dots, x_n)$$

expressing that the values of x_n depend functionally on the values of x_1, \dots, x_n . Inclusion logic, instead, extends first-order logic with inclusion atoms

$$\mathbf{x} \subseteq \mathbf{y}$$

which express that the set of values of \mathbf{x} is included in the set of the values of \mathbf{y} . We study the expressive power of two syntactic fragments of inclusion logic under the lax team semantics. These two fragments, $\text{FO}(\subseteq)(k\forall)$ and $\text{FO}(\subseteq)(k\text{-inc})$, are defined by restricting the number of universal quantifiers or the arity of inclusion atom to k , respectively. We will show that $\text{FO}(\subseteq)(k\forall)$ captures $\text{FO}(\subseteq)$ already with $k = 1$ and that the fragments $\text{FO}(\subseteq)(k\text{-inc})$ give rise to an infinite expressivity hierarchy.

Since the introduction of dependence logic in 2007, many interesting variants of it have been introduced. One reason for this orientation is the semantical framework that is being used. Team semantics, introduced by Hodges in 1997 [12], provides a natural way to extend first-order logic with many different kind of dependency notions. Although many of these notions have been extensively studied in database theory since the 70s, with team semantics the novelty comes

* The author was supported by grant 264917 of the Academy of Finland.

** Department of Mathematics and Statistics, University of Helsinki, Finland.
miika.hannula@helsinki.fi

from the fact that also interpretations for logical connectives and quantifiers are provided.

In expressive power $\text{FO}(=(\dots))$ is equivalent to existential second-order logic (ESO) [16]. For some variants of $\text{FO}(=(\dots))$, the correspondence to ESO does not hold or it can depend on which version of team semantics is being used. For instance, $\text{FO}(\subseteq)$ corresponds in expressive power to ESO if we use the so-called strict team semantics [5]. Under the lax team semantics, $\text{FO}(\subseteq)$ corresponds to greatest fixed point logic (GFP) [6] which captures PTIME over finite ordered models. In $\text{FO}(=(\dots))$ no separation between the strict and the lax version of team semantics exists since dependence atoms satisfy the so-called downward closure property. In the following we briefly list some complexity theoretical aspects of $\text{FO}(=(\dots))$ and its variants.

- $\text{FO}(=(\dots))$ extended with the so-called intuitionistic implication \rightarrow (introduced in [1]) increases the expressive power of $\text{FO}(=(\dots))$ to full second-order logic [17].
- The model checking problem of $\text{FO}(=(\dots))$, and many of its variants, was recently shown to be NEXPTIME-complete. Moreover, for any variant of $\text{FO}(=(\dots))$ whose atoms are PTIME-computable, the corresponding model checking problem is contained in NEXPTIME [8].
- The non-classical interpretation of disjunction in $\text{FO}(=(\dots))$ has the effect that the model checking problems of $\phi_1 := =(x, y) \vee =(u, v)$ and $\phi_2 := =(x, y) \vee =(u, v) \vee =(u, v)$ are already NL-complete and NP-complete, respectively [15].

This article pursues the line of study taken in [3] and [5] where syntactical fragments of dependence and independence logic ($\text{FO}(\perp_c)$) were investigated, respectively. $\text{FO}(\perp_c)$ extends first-order logic by conditional independence atoms

$$\mathbf{y} \perp_{\mathbf{x}} \mathbf{z}$$

with the informal meaning that the values of \mathbf{y} and \mathbf{z} are independent of each other, given any value of \mathbf{x} . As $\text{FO}(\subseteq)$, also $\text{FO}(\perp_c)$ does not have downward closure and is sensitive to the choice between the lax and the strict version of team semantics. For a sequence of atoms \mathcal{C} , we use $\text{FO}(\mathcal{C})$ to denote the logic obtained by adding the atoms listed in \mathcal{C} to first-order logic. $\text{FO}(\mathcal{C})(k\forall)$ denotes the sentences of $\text{FO}(\mathcal{C})$ in which at most k variables are universally quantified. In [3] it was shown that

$$\text{FO}(=(\dots))(k\forall) \leq \text{ESO}_f(k\forall) \leq \text{FO}(=(\dots))(2k\forall)$$

where $\text{ESO}_f(k\forall)$ denotes the skolem normal form ESO sentences in which at most k universally quantified first-order variables appear. In [5] it was shown that (under the lax team semantics)

- $\text{FO}(\perp)(2\forall) = \text{FO}(\perp)$ and
- $\text{FO}(\perp, \subseteq)(1\forall) = \text{FO}(\perp, \subseteq)$

where $\text{FO}(\perp)$ is the sublogic of $\text{FO}(\perp_c)$ allowing only so-called pure independence atoms $\mathbf{x} \perp \mathbf{y}$. Moreover, it is known that $\text{FO}(\perp)$ is equivalent in expressive power to $\text{FO}(\perp, \subseteq)$ and $\text{FO}(\perp_c)$ [4, 7].

Also arity fragments of $\text{FO}(\mathcal{C})$ were defined. By $\text{FO}(\mathcal{C})(k\text{-dep})$ we denote the sentences of $\text{FO}(\mathcal{C})$ in which dependence atoms of the form $\mathbf{=}(x_1, \dots, x_{n+1})$ with $n \leq k$ may appear. $\text{FO}(\mathcal{C})(k\text{-ind})$ denotes the sentences of $\text{FO}(\perp_c)$ in which independence atoms containing at most $k + 1$ different variables may appear. It was shown in [3, 5] that (under the lax team semantics)

$$\text{ESO}(k\text{-ary}) = \text{FO}(\mathbf{=})(k\text{-dep}) = \text{FO}(\perp_c)(k\text{-ind})$$

where $\text{ESO}(k\text{-ary})$ denote the sentences of ESO in which the quantified functions and relations have arity at most k . This yields an infinite arity hierarchy for both $\text{FO}(\mathbf{=})(k\text{-dep})$ and $\text{FO}(\perp_c)$ since the property "R is even" is definable in $\text{ESO}(k\text{-ary})$ but not in $\text{ESO}(k - 1\text{-ary})$, for k -ary R [2].

The main contribution of this article is to show that arity fragments of inclusion logic also give rise to an infinite expressivity hierarchy. We let $\text{FO}(\mathcal{C})(k\text{-inc})$ denote the $\text{FO}(\mathcal{C})$ sentences in which at most k -ary inclusion atoms (i.e. atoms of the form $\mathbf{x} \subseteq \mathbf{y}$ where $|\mathbf{x}| = |\mathbf{y}| \leq k$) may appear. For proving the claim, we define, for each $k \geq 2$, a graph property which is definable in $\text{FO}(\subseteq)(k\text{-inc})$ but not in $\text{FO}(\subseteq)(k - 1\text{-inc})$. The non-definability part of the proof will be based on Martin Grohe's work in fixed point logics in [10] where analogous results for TC, LFP, IFP and PFP were proved. We will also give a negative answer to the open question presented in [5]; that was, whether the fragments $\text{FO}(\subseteq)(k\forall)$ give rise to an infinite expressivity hierarchy. This will be done by showing that $\text{FO}(\subseteq)(1\forall) = \text{FO}(\subseteq)$. However, if the strict version of team semantics is used, then we obtain $\text{FO}(\subseteq)(k\forall) < \text{FO}(\subseteq)(k + 1\forall)$ [11].

2 Preliminaries

In this section we give a short introduction to dependence, independence and inclusion logic.

2.1 Notation

Unless otherwise stated, we use x_1, x_2, \dots to denote variables and t_1, t_2, \dots to denote terms. Analogously, bolded versions $\mathbf{x}_1, \mathbf{x}_2, \dots$ and $\mathbf{t}_1, \mathbf{t}_2, \dots$ are used to denote tuples of variables and tuples of terms, respectively. For tuples \mathbf{a} and \mathbf{b} , we write \mathbf{ab} for the concatenation of the tuples. If f is a unary function and (x_1, \dots, x_n) is a sequence listing members of $\text{Dom}(f)$, then we write $f(x_1, \dots, x_n)$ for $(f(x_1), \dots, f(x_n))$.

2.2 Inclusion logic

The syntax of $\text{FO}(\subseteq)$ is obtained by adding inclusion atoms to the syntax of first-order logic.

Definition 1. $\text{FO}(\subseteq)$ is defined by the following grammars. Note that in an inclusion atom $\mathbf{x}_1 \subseteq \mathbf{x}_2$, the tuples \mathbf{x}_1 and \mathbf{x}_2 must be of the same length.

$$\phi ::= \mathbf{x}_1 \subseteq \mathbf{x}_2 \mid t_1 = t_2 \mid \neg t_1 = t_2 \mid R(\mathbf{t}) \mid \neg R(\mathbf{t}) \mid (\phi \vee \psi) \mid (\phi \wedge \psi) \mid \forall x \phi \mid \exists x \phi.$$

$\text{FO}(=)$ and $\text{FO}(\perp_c)$ are obtained from Definition 1 by replacing inclusion atoms $\mathbf{x}_1 \subseteq \mathbf{x}_2$ with dependence atoms $=(\mathbf{x}_1, \mathbf{x}_2)$ and conditional independence atoms $\mathbf{x}_2 \perp_{\mathbf{x}_1} \mathbf{x}_3$, respectively. Pure independence logic $\text{FO}(\perp)$ is a fragment of $\text{FO}(\perp_c)$ where only pure independence atoms $\mathbf{x}_1 \perp \mathbf{x}_2$ (i.e. atoms of the form $\mathbf{x}_1 \perp_{\emptyset} \mathbf{x}_2$) may appear. Also, for any sequence \mathcal{C} of dependency atoms of $\{\subseteq, =, \perp_c, \perp\}$ we use $\text{FO}(\mathcal{C})$ to denote the logic obtained from Definition 1 by replacing inclusion atoms with atoms listed in \mathcal{C} .

In order to define semantics for these logics, we need to define the concept of a *team*. Let \mathfrak{M} be a model with the domain M . We assume that all our models have at least two elements.¹ An *assignment* over \mathfrak{M} is a finite function that maps variables to elements of M . A *team* X of M with the domain $\text{Dom}(X) = \{x_1, \dots, x_n\}$ is a set of assignments from $\text{Dom}(X)$ into M . If X is a team of M and $F : X \rightarrow \mathcal{P}(M) \setminus \{\emptyset\}$, then we use $X[F/x]$ to denote the team $\{s(a/x) \mid s \in X, a \in F(s)\}$ and $X[M/x]$ for $\{s(a/x) \mid s \in X, a \in M\}$. Also one should note that if s is an assignment, then $\mathfrak{M} \models_s \phi$ refers to Tarski semantics and $\mathfrak{M} \models_{\{s\}} \phi$ refers to team semantics.

Definition 2. For a model \mathfrak{M} , a team X and a formula in $\text{FO}(\subseteq, =, \perp_c)$, the satisfaction relation $\mathfrak{M} \models_X \phi$ is defined as follows:

- $\mathfrak{M} \models_X \alpha \Leftrightarrow \forall s \in X (\mathfrak{M} \models_s \alpha)$, when α is a first-order literal,
- $\mathfrak{M} \models_X \mathbf{x}_1 \subseteq \mathbf{x}_2 \Leftrightarrow \forall s \in X \exists s' \in X (s(\mathbf{x}_1) = s'(\mathbf{x}_2))$,
- $\mathfrak{M} \models_X \mathbf{x}_2 \perp_{\mathbf{x}_1} \mathbf{x}_3 \Leftrightarrow \forall s, s' \in X (s(\mathbf{x}_1) = s'(\mathbf{x}_2) \Rightarrow \exists s'' \in X (s''(\mathbf{x}_1) = s(\mathbf{x}_1), s''(\mathbf{x}_2) = s'(\mathbf{x}_2), s''(\mathbf{x}_3) = s'(\mathbf{x}_3)))$,
- $\mathfrak{M} \models_X =(\mathbf{x}_1, \mathbf{x}_2) \Leftrightarrow \forall s, s' \in X (s(\mathbf{x}_1) = s'(\mathbf{x}_1) \Rightarrow s(\mathbf{x}_2) = s'(\mathbf{x}_2))$,
- $\mathfrak{M} \models_X \phi \wedge \psi \Leftrightarrow \mathfrak{M} \models_X \phi$ and $\mathfrak{M} \models_X \psi$,
- $\mathfrak{M} \models_X \phi \vee \psi \Leftrightarrow \mathfrak{M} \models_Y \phi$ and $\mathfrak{M} \models_Z \psi$, for some $Y \cup Z = X$,
- $\mathfrak{M} \models_X \exists x \phi \Leftrightarrow \mathfrak{M} \models_{X[F/x]} \phi$, for some $F : X \rightarrow \mathcal{P}(M) \setminus \{\emptyset\}$,
- $\mathfrak{M} \models_X \forall x \phi \Leftrightarrow \mathfrak{M} \models_{X[M/x]} \phi$.

If $\mathfrak{M} \models_X \phi$, then we say that X satisfies ϕ in \mathfrak{M} . If ϕ is a sentence and $\mathfrak{M} \models_{\{\emptyset\}} \phi$, then we say that ϕ is true in \mathfrak{M} , and write $\mathfrak{M} \models \phi$.

Note that in Definition 2, we obtain the *lax* version of team semantics. The *strict* version of team semantics is defined otherwise as in Definition 2 except that only disjoint subteams are allowed to witness split disjunction and existential quantification ranges over M instead of non-empty subsets of M . (See [4] for more information.)

First-order formulae are *flat* in the following sense (the proof is a straightforward structural induction).

¹ This assumption is needed in Theorem 3.

² $\{\emptyset\}$ denotes the team that consists of the empty assignment.

Theorem 1 (Flatness). *For a model \mathfrak{M} , a first-order formula ϕ and a team X , the following are equivalent:*

- $\mathfrak{M} \models_X \phi$,
- $\mathfrak{M} \models_{\{s\}} \phi$ for all $s \in X$,
- $\mathfrak{M} \models_s \phi$ for all $s \in X$.

By $\text{Fr}(\phi)$ we denote the set of variables that appear free in ϕ . If X is a team and V a set of variables, then $X \upharpoonright V$ denotes the team $\{s \upharpoonright V \mid s \in X\}$. Now, all formulae satisfy the following *locality* property. Note that this is not true under the strict team semantics.

Theorem 2 (Locality). *Let \mathfrak{M} be a model, X be a team, $\phi \in \text{FO}(\subseteq, =(\dots), \perp_c)$ and V a set of variables such that $\text{Fr}(\phi) \subseteq V \subseteq \text{Dom}(X)$. Then*

$$\mathfrak{M} \models_X \phi \Leftrightarrow \mathfrak{M} \models_{X \upharpoonright V} \phi.$$

We say that formulae $\phi, \psi \in \text{FO}(\subseteq, =(\dots), \perp_c)$ are *logically equivalent*, written $\phi \equiv \psi$, if for all models \mathfrak{M} and teams X such that $\text{Fr}(\phi) \cup \text{Fr}(\psi) \subseteq \text{Dom}(X)$,

$$\mathfrak{M} \models_X \phi \Leftrightarrow \mathfrak{M} \models_X \psi.$$

We obtain the following normal form theorem.

Theorem 3 ([5]). *Any formula $\phi \in \text{FO}(\subseteq, =(\dots), \perp_c)$ is logically equivalent to a formula ϕ' such that*

- ϕ' is of the form $Q^1 x_1 \dots Q^n x_n \psi$ where ψ is quantifier-free,
- any literal or dependency atom which occurs in ϕ' occurred already in ϕ ,
- the number of universal quantifiers in ϕ' is the same as the number of universal quantifiers in ϕ .

For logics \mathcal{L} and \mathcal{L}' , we write $\mathcal{L} \leq \mathcal{L}'$, if for every signature τ , every $\mathcal{L}[\tau]$ -sentence is logically equivalent to some $\mathcal{L}'[\tau]$ -sentence. We write $\mathcal{L} \leq_{\mathcal{O}} \mathcal{L}'$ if $\mathcal{L} \leq \mathcal{L}'$ is true in finite linearly ordered models. Equality and inequality relations are obtained from \leq naturally. We end this section with the following list of theorems characterizing the expressive powers of our logics.

Theorem 4 ([16, 9, 7, 6]).

- $\text{FO}(=(\dots)) = \text{FO}(\perp_c) = \text{FO}(\perp) = \text{ESO}$,
- $\text{FO}(\subseteq) = \text{GFP}$.

3 Hierarchies in Inclusion Logic

In this section we consider universal and arity fragments of inclusion logic. In Subsection 3.1 we will define these fragments and also concepts of strictness and collapse of a hierarchy. In Subsection 3.2 and 3.3 we will prove collapse of the universal hierarchy and strictness of the arity hierarchy, respectively.

3.1 Syntactical Fragments

Definition 3. Let \mathcal{C} be a list of dependencies of $\{\subseteq, =(\dots), \perp_c, \perp\}$. Then universal and arity fragments of $\text{FO}(\mathcal{C})$ are defined as follows:

- $\text{FO}(\mathcal{C})(k\forall)$ is the class of $\text{FO}(\mathcal{C})$ formulae in which at most k universal quantifiers may appear,
- $\text{FO}(\mathcal{C})(k\text{-inc})$ is the class of $\text{FO}(\mathcal{C})$ formulae in which inclusion atoms of the form $\mathbf{x}_1 \subseteq \mathbf{x}_2$ where \mathbf{x}_1 and \mathbf{x}_2 are sequences of length at most k , may appear,
- $\text{FO}(\mathcal{C})(k\text{-dep})$ is the class of $\text{FO}(\mathcal{C})$ formulae in which dependence atoms of the form $=(\mathbf{x}_1, \mathbf{x}_2)$ where $\mathbf{x}_1\mathbf{x}_2$ is a sequence of length at most $k + 1$, may appear,
- $\text{FO}(\mathcal{C})(k\text{-ind})$ is the class of $\text{FO}(\mathcal{C})$ formulae in which conditional independence atoms of the form $\mathbf{x}_2 \perp_{\mathbf{x}_1} \mathbf{x}_3$ where $\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$ is a sequence listing at most $k + 1$ distinct variables, may appear.

For a sequence of logics $(\mathcal{L}_k)_{k \in \mathbb{N}}$, we say that the \mathcal{L}_k -hierarchy collapses at level m if $\mathcal{L}_m = \bigcup_{k \in \mathbb{N}} \mathcal{L}_k$. If the hierarchy does not collapse at any level, then we say that it is strict.

As mentioned before, we will show that the $\text{FO}(\subseteq)(k\forall)$ -hierarchy collapses already at level 1 but $\text{FO}(\subseteq)(k\text{-inc})$ forms a strict hierarchy which holds already in finite models.

3.2 Collapse of the Universal Hierarchy

We will first show that the universal hierarchy of inclusion logic collapses. This is done by introducing a translation where all universal quantifiers are removed, and new existential quantifiers, new inclusion atoms and one new universal quantifier are added. The translation will hold already at the level of formulae.

Theorem 5. $\text{FO}(\subseteq)(1\forall) = \text{FO}(\subseteq)$.

Proof. Let $\phi \in \text{FO}(\subseteq)$ be a formula. We will define a $\phi' \in \text{FO}(\subseteq)(1\forall)$ such that $\phi \equiv \phi'$. By Theorem 3 we may assume that ϕ is of the form

$$Q^1 x_1 \dots Q^n x_n \theta$$

where θ is quantifier-free. We let

$$\phi' := \exists x_1 \dots \exists x_n \forall y \left(\bigwedge_{\substack{1 \leq i \leq n \\ Q^i = \forall}} z x_1 \dots x_{i-1} y \subseteq z x_1 \dots x_{i-1} x_i \wedge \theta \right)$$

where z lists $\text{Fr}(\phi)$. Let now \mathfrak{M} be a model and X a team such that $\text{Fr}(\phi) \subseteq \text{Dom}(X)$; we show that $\mathfrak{M} \models_X \phi \Leftrightarrow \mathfrak{M} \models_X \phi'$. By Theorem 2 we may assume without loss of generality that $\text{Fr}(\phi) = \text{Dom}(X)$. Assume first that $\mathfrak{M} \models_X \phi$

when there are, for $1 \leq i \leq n$, functions $F_i : X[F_1/x_1] \dots [F_{i-1}/x_{i-1}] \rightarrow \mathcal{P}(M) \setminus \{\emptyset\}$ such that $F_i(s) = M$ if $Q^i = \forall$, and $\mathfrak{M} \models_{X'} \theta$ where

$$X' := X[F_1/x_1] \dots [F_n/x_n].$$

For $\mathfrak{M} \models_X \phi'$, it suffices to show that

$$\mathfrak{M} \models_{X'[M/y]} \bigwedge_{\substack{1 \leq i \leq n \\ Q^i = \forall}} \mathbf{z}x_1 \dots x_{i-1}y \subseteq \mathbf{z}x_1 \dots x_{i-1}x_i \wedge \theta. \quad (1)$$

By Theorem 2 $\mathfrak{M} \models_{X'[M/y]} \theta$, so it suffices to consider only the new inclusion atoms of (1). So let $1 \leq i \leq n$ be such that $Q^i = \forall$ and let $s \in X'[M/y]$; we need to find a $s' \in X'[M/y]$ such that $s(\mathbf{z}x_1 \dots x_{i-1}y) = s'(\mathbf{z}x_1 \dots x_{i-1}x_i)$. Now, since $Q^i = \forall$, we note that $s(s(y)/x_i) \in X' \upharpoonright (\text{Fr}(\phi) \cup \{x_1, \dots, x_i\})$. Therefore we may choose s' to be any extension of $s(s(y)/x_i)$ in $X'[M/y]$.

For the other direction, assume that $\mathfrak{M} \models_X \phi'$. Then for $1 \leq i \leq n$, there are functions $F_i : X[F_1/x_1] \dots [F_{i-1}/x_{i-1}] \rightarrow \mathcal{P}(M) \setminus \{\emptyset\}$ such that (1) holds, for $X' := X[F_1/x_1] \dots [F_n/x_n]$. By Theorem 2 $\mathfrak{M} \models_{X'} \theta$, so it suffices to show that, for all $1 \leq i \leq n$ with $Q^i = \forall$, F_i is the constant function which maps assignments to M . So let i be of the above kind, and let $s \in X[F_1/x_1] \dots [F_{i-1}/x_{i-1}]$ and $a \in M$. We need show that $s(a/x_i) \in X[F_1/x_1] \dots [F_i/x_i]$. First note that since y is universally quantified, $s(a/y)$ has an extension s_0 in $X'[M/y]$. Therefore, by (1), there is $s_1 \in X'[M/y]$ such that $s_0(\mathbf{z}x_1 \dots x_{i-1}y) = s_1(\mathbf{z}x_1 \dots x_{i-1}x_i)$. Since now s_1 agrees with s in $\text{Fr}(\phi) \cup \{x_1, \dots, x_{i-1}\}$ and maps x_i to a , we obtain that

$$s(a/x_i) = s_1 \upharpoonright (\text{Fr}(\phi) \cup \{x_1, \dots, x_i\}) \in X[F_1/x_1] \dots [F_i/x_i].$$

□

3.3 Strictness of the Arity Hierarchy

In this section we will show that the following strict arity hierarchy holds (already in finite models).

Theorem 6. For $k \geq 2$, $\text{FO}(\subseteq)(k-1\text{-inc}) < \text{FO}(\subseteq)(k\text{-inc})$.

For proving this, we will use the earlier work of Grohe in [10] where an analogous result was proved for TC, LFP, IFP and PFP. More precisely, it was shown that, for $k \geq 2$,

$$\text{TC}^k \not\leq \text{PFP}^{k-1} \quad (2)$$

where the superscript part gives the maximum arity allowed for the fixed point operator. Since $\text{TC}^k \leq \text{LFP}^k \leq \text{IFP}^k \leq \text{PFP}^k$, a strict arity hierarchy is obtained for each of these logics.

We start by fixing τ as the signature consisting of one binary relation symbol E and $2k$ constant symbols $b_1, \dots, b_k, c_1, \dots, c_k$. The idea is to present a $\text{FO}(\subseteq)(k\text{-inc})[\tau]$ -definable graph property, and show that it is not definable in $\text{FO}(\subseteq)$

$(k-1\text{-inc})[\tau]$. This graph property will actually be negated version of the one that separates the fragments in (2). For this, we first define a first-order formula indicating that the k -tuples \mathbf{x} and \mathbf{y} form a $2k$ -clique in a graph. Namely, we define $\text{EDGE}_k(\mathbf{x}, \mathbf{y})$ as follows:

$$\text{EDGE}_k(\mathbf{x}, \mathbf{y}) := \bigwedge_{1 \leq i, j \leq k} E(x_i, y_j) \wedge \bigwedge_{1 \leq i \neq j \leq k} (E(x_i, x_j) \wedge E(y_i, y_j)).$$

Then the non-trivial part is to show that negation of the transitive closure formula $[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ is not definable in $\text{FO}(\subseteq)(k-1\text{-inc})[\tau]$. It is definable in $\text{FO}(\subseteq)(k\text{-inc})[\tau]$ by the following theorem.

Theorem 7 ([4]). *Let $\psi(\mathbf{x}, \mathbf{y})$ be any first-order formula, where \mathbf{x} and \mathbf{y} are tuples of disjoint variables of the same arity. Furthermore, let $\psi'(\mathbf{x}, \mathbf{y})$ be the result of writing $\neg\psi(\mathbf{x}, \mathbf{y})$ in negation normal form. Then, for all suitable models \mathfrak{M} and all suitable pairs \mathbf{b}, \mathbf{c} of constant term tuples of the model,*

$$\mathfrak{M} \models \phi \Leftrightarrow \mathfrak{M} \models \neg[TC_{\mathbf{x}, \mathbf{y}} \psi](\mathbf{b}, \mathbf{c}),$$

for ϕ defined as

$$\exists z(\mathbf{b} \subseteq z \wedge z \neq \mathbf{c} \wedge \forall w(\psi'(z, w) \vee w \subseteq z)).$$

Note that ϕ is not yet of the right form since Definition 1 does not allow terms to appear in inclusion atoms. This is however not a problem since we can replace all terms that appear in inclusion atoms with new existentially quantified variables.

Hence, for Theorem 6, it suffices to prove that $\neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ is not definable in $\text{FO}(\subseteq)(k-1\text{-inc})[\tau]$. In this part we will follow the work in [10]. We first define a simple structure $\mathfrak{A}(k, n)$, for $k, n \geq 1$. $\mathfrak{A}(k, n)$ consists of two disjoint E_k -paths of length n i.e.

$$\begin{aligned} A &:= \{1, \dots, n\} \times \{-k, \dots, -1, 1, \dots, k\} \\ E_k^{\mathfrak{A}} &:= \{(I, -1), \dots, (I, -k)(I+1, -1), \dots, (I+1, -k) \mid 1 \leq I \leq n\} \\ &\quad \cup \{(I, 1), \dots, (I, k)(I+1, 1), \dots, (I+1, k) \mid 1 \leq I \leq n\}. \end{aligned}$$

The following theorem generates a graph \mathfrak{A} of the form $\mathfrak{A}(k, n)$, for $E_k^{\mathfrak{A}} := \text{EDGE}_k^{\mathfrak{A}}$, with many useful properties. It was originally proved by Grohe using a method of Hrushovski [13] to extend partial isomorphisms of finite graphs.

Theorem 8 ([10]). *Let $k, n \geq 2$. Then there exists a graph $\mathfrak{A} = \mathfrak{A}(k, n)$ such that:*

1. *There exists a mapping $\text{row}: A \rightarrow \{1, \dots, n\}$ such that*

$$\forall a, b \in A : (E^{\mathfrak{A}} ab \Rightarrow \text{row}(b) - \text{row}(a) \leq 1).$$

2. *There exists an automorphism ε of \mathfrak{A} that is self-inverse and preserves the rows i.e.*

- $\varepsilon^{-1} = \varepsilon$,
 - $\forall a \in A : \text{row}(\varepsilon(a)) = \text{row}(a)$.
3. There exist tuples $\mathbf{b}, \mathbf{c} \in A^k$ in the first and last row respectively (i.e. $\forall i \leq k : (\text{row}(b_i) = 1 \wedge \text{row}(c_i) = n)$) such that

$$\mathfrak{A} \models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c}) \text{ and } \mathfrak{A} \not\models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \varepsilon(\mathbf{c})).^3$$

4. For all $a_1, \dots, a_{k-1} \in A$ there exists an automorphism f that is self-inverse, preserves the rows, and maps a_1, \dots, a_{k-1} according to ε , but leaves all elements in rows of distance > 1 from $\text{row}(a_1), \dots, \text{row}(a_{k-1})$ fixed i.e.
- $f^{-1} = f$
 - $\forall a \in A : \text{row}(f(a)) = \text{row}(a)$,
 - $\forall i \leq k-1 : f(a_i) = \varepsilon(a_i)$,
 - for each $a \in A$ with $\forall i \leq k-1 : |\text{row}(a) - \text{row}(a_i)| > 1$ we have $f(a) = a$.

Using this theorem we will prove the following lemma.

Lemma 1. *Let $k \geq 2$ and let τ be a signature consisting of a binary relation symbol E and $2k$ constant symbols $b_1, \dots, b_k, c_1, \dots, c_k$. Then $\neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ is not definable in $\text{FO}(\subseteq)(k-1\text{-inc})[\tau]$.*

The outline of the proof is listed below:

1. First we assume to the contrary that there is a $\phi(\mathbf{b}, \mathbf{c}) \in \text{FO}(\subseteq)(k-1\text{-inc})[\tau]$ which is equivalent to $\neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$.
2. By Theorem 3 we may assume that ϕ is of the form $Q^1 x_1 \dots Q^m x_m \theta$ where θ is a quantifier-free $\text{FO}(\subseteq)(k-1\text{-inc})[\tau]$ formula.
3. We let $n = 2^{m+2}$ and obtain a graph \mathfrak{A} for which items 1-4 of Theorem 8 hold, for k, n . In particular, we find tuples \mathbf{b} and \mathbf{c} such that $\mathfrak{A} \models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ and $\mathfrak{A} \not\models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \varepsilon(\mathbf{c}))$. Then by the counter assumption $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models \phi$ when we find, for $1 \leq i \leq m$, functions

$$F_i : \{\emptyset\}[F_1/x_1] \dots [F_{i-1}/F_{i-1}] \rightarrow \mathcal{P}(A) \setminus \{\emptyset\}$$

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

$$(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_X \theta$$

where $X := \{\emptyset\}[F_1/x_1] \dots [F_m/x_m]$.

4. From X we will construct a team X^* such that

$$(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{X^*} \theta. \tag{3}$$

For this, we define operations auto and swap for teams Y of A with $\text{Dom}(Y) = \{x_1, \dots, x_m\}$. We first let $\text{auto}(Y)$ be the set of assignments $f \circ s$ where $s \in Y$ and f is a composition of automorphisms of \mathfrak{A} that swap $k-1$ -tuples of A but

³ In [10], \mathbf{c} and $\varepsilon(\mathbf{c})$ are here placed the other way round. This is however not a problem since ε is self-inverse and preserves the rows.

leave elements in tuples \mathbf{b} and \mathbf{c} fixed. Then we let $\text{swap}(Y)$ transform each assignment s of Y to a s' such that s' corresponds to Duplicator's choices in a single play of $\text{EF}_m((\mathfrak{A}, \mathbf{b}, \mathbf{c}), (\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})))$ where Spoiler picks members of $(\mathfrak{A}, \mathbf{b}, \mathbf{c})$ according to s and Duplicator picks members of $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c}))$ according to her winning strategy.⁴ We then let $X^* := \text{swap}(\text{auto}(X))$ and show (3).

5. At last, we will show that X^* can be constructed by quantifying $Q^1 x_1 \dots Q^m x_m$ in \mathfrak{A} over $\{\emptyset\}$.

Hence we will obtain that $\mathfrak{A} \models \phi(\mathbf{b}, \varepsilon(\mathbf{c}))$. But now, since $\mathfrak{A} \not\models \neg[\text{TC}_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \varepsilon(\mathbf{c}))$, this contradicts with the assumption that $\phi(\mathbf{b}, \mathbf{c})$ defines $\neg[\text{TC}_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$.

Let us now proceed to the proof.

Proof (Lemma 1). We may start from item 4 of the previous list. Hence we have

$$(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_X \theta, \quad (4)$$

for $X := \{\emptyset\}[F_1/x_1] \dots [F_m/x_m]$, and the first step is to construct a team X^* such that

$$(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{X^*} \theta. \quad (5)$$

For this, we first define the operation auto . By item 4 of Theorem 8, for all \mathbf{a} listing $a_1, \dots, a_{k-1} \in A$ there exists an automorphism $f_{\mathbf{a}}$ which maps \mathbf{a} pointwise to $\varepsilon(\mathbf{a})$, but leaves all elements in rows of distance > 1 from $\text{row}(a_1), \dots, \text{row}(a_{k-1})$ fixed. Let $\mathcal{F} = (F, \circ)$ be the group generated by the automorphisms $f_{\mathbf{a}}$ where $f_{\mathbf{a}}$ is obtained from item 4 of Theorem 8 and \mathbf{a} is a sequence listing $a_1, \dots, a_{k-1} \in A$ such that $2 < \text{row}(a_i) < n - 1$, for $1 \leq i \leq k - 1$. For a team Y of A , we then let

$$\text{auto}(Y) := \{f \circ s \mid f \in \mathcal{F}, s \in Y\}.$$

Next we will define the operation swap . For this, we will first define mappings mid and h . We let mid map m -sequences of $\{1, \dots, n\}$ into $\{1, \dots, n\}$ so that, for any $\mathbf{p} := (p_1, \dots, p_m)$ and $\mathbf{q} := (q_1, \dots, q_m)$ in $\{1, \dots, n\}^m$,

1. $1 < \text{mid}(\mathbf{p}) < n$,
2. $\forall i \leq m : \text{mid}(\mathbf{p}) \neq p_i$,
3. $\forall l \leq n$: if $\mathbf{p} \upharpoonright \{1, \dots, l\} = \mathbf{q} \upharpoonright \{1, \dots, l\}$, then $\forall i \leq l : p_i < \text{mid}(\mathbf{p})$ iff $q_i < \text{mid}(\mathbf{q})$.

⁴ A play of this kind is illustrated in Fig. 1. The idea is that after each round $i \leq m$, M and N are placed so that $N - M \geq 2^{m+1-i}$. Also for each $j \leq i$, $y_j = x_j$ if $\text{row}(x_j) \leq M$, and $y_j = \varepsilon(x_j)$ if $\text{row}(x_j) \geq N$, where y_j and x_j represent Duplicator's and Spoiler's choices, respectively. In the picture, α and β represent two alternative choices Spoiler can make at the fourth round. If Spoiler chooses $x_4 := \alpha$, then Duplicator chooses $y_4 := \alpha$, and M is moved to $\text{row}(\alpha)$. If Spoiler chooses $x_4 := \beta$, then Duplicator chooses $y_4 := \varepsilon(\beta)$, and N is moved to $\text{row}(\beta)$. Proceeding in this way we obtain that at the final stage m , $(\mathfrak{A}, \mathbf{b}, \mathbf{c}, x_1, \dots, x_m)$ and $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c}), y_1, \dots, y_m)$ agree on all atomic $\text{FO}[\tau]$ formulae.

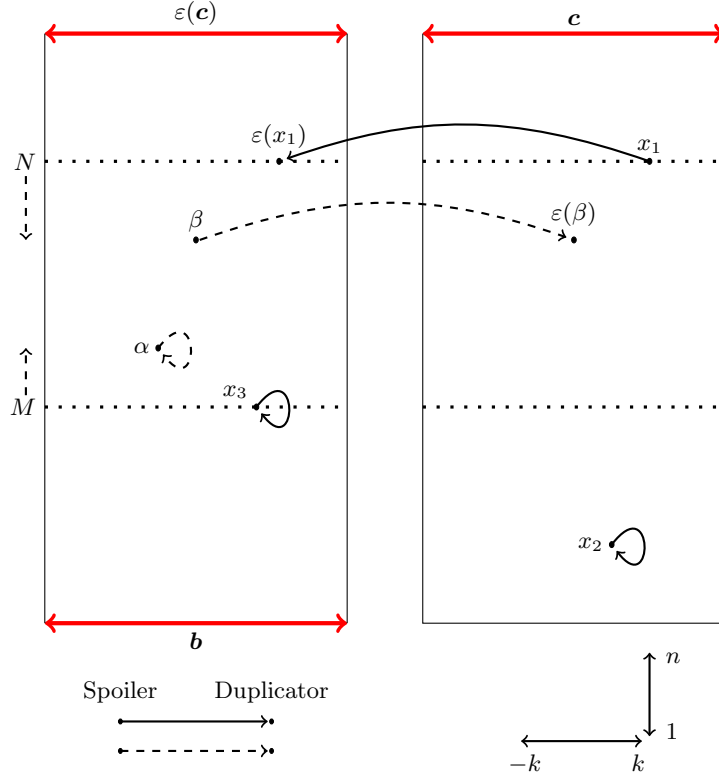


Fig. 1

This can be done by following the strategy illustrated in Fig. 1. We shall explain this in detail in the following. Let $\mathbf{p} := (p_1, \dots, p_m)$ be a sequence listing natural numbers of $\{1, \dots, n\}$. For $\text{mid}(\mathbf{p})$, we first show how to choose M and N , for each $0 \leq i \leq m$, so that

- $N - M \geq 2^{m+1-i}$,
- $\forall j \leq i : p_j \leq M$ or $p_j \geq N$.

We do this inductively as follows. We let $M := 1$ and $N := n$, for $i = 0$. Since $n = 2^{m+2}$, clearly the conditions above hold. Assume that M and N are defined for i so that the conditions above hold; we define M' and N' for $i + 1$ as follows:

1. If $p_{i+1} - M \leq N - p_{i+1}$, then we let $M' := \max\{M, p_{i+1}\}$ and $N' := N$.
2. If $p_{i+1} - M > N - p_{i+1}$, then we let $M' := M$ and $N' := \min\{N, p_{i+1}\}$.

Note that in both cases $\forall j \leq i + 1 : p_j \leq M'$ or $p_j \geq N'$, and

$$N' - M' \geq \left\lceil \frac{N - M}{2} \right\rceil \geq 2^{m+1-(i+1)}.$$

Proceeding in this way we conclude that at the final stage m we have $N - M \geq 2$ with no p_1, \dots, p_m strictly in between M and N . We then choose $\text{mid}(\mathbf{p})$ as any number in $]M, N[$. Note that defining mid in this way we are able to meet the conditions 1-3.

After this we define a mapping $h : \{x_1, \dots, x_m\}A \rightarrow \{x_1, \dots, x_m\}A$. For an assignment $s : \{x_1, \dots, x_m\} \rightarrow A$, the assignment $h(s) : \{x_1, \dots, x_m\} \rightarrow A$ is defined as follows:

$$h(s)(x_i) = \begin{cases} s(x_i) & \text{if } \text{row}(s(x_i)) < \text{mid}(\text{row}(s(\mathbf{x}))), \\ \varepsilon \circ s(x_i) & \text{if } \text{row}(s(x_i)) > \text{mid}(\text{row}(s(\mathbf{x}))), \end{cases}$$

where $\mathbf{x} := (x_1, \dots, x_m)$. For a team Z of A with $\text{Dom}(Z) = \{x_1, \dots, x_m\}$, we now let

$$\text{swap}(Z) := \{h(s) \mid s \in Z\},$$

and define, for each $Y \subseteq X$,

$$Y^* := \text{swap}(\text{auto}(Y)).$$

With X^* now defined, we will next show that (5) holds. Without loss of generality we may assume that if a constant symbol b_j (or c_j) appears in an atomic subformula α of θ , then α is of the form $x_i = b_j$ (or $x_i = c_j$) where x_i is an existentially quantified variable of the quantifier prefix. Hence and by (4), it now suffices to show that for all $Y \subseteq X$ and all quantifier-free $\psi \in \text{FO}(\subseteq)(k-1\text{-inc})[\tau]$ with the above restriction for constants,

$$(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_Y \psi \Rightarrow (\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{Y^*} \psi.$$

This can be done by induction on the complexity of the quantifier-free ψ . Since $Y^* \cup Z^* = (Y \cup Z)^*$, for $Y, Z \subseteq X$, it suffices to consider only the case where ψ is an atomic or negated atomic formula. For this, assume that $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_Y \psi$; we will show that

$$(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{Y^*} \psi. \tag{6}$$

Now ψ is either of the form $x_i = b_j$, $x_i = c_j$, $x_i = x_j$, $\neg x_i = x_j$, $E(x_i, x_j)$, $\neg E(x_i, x_j)$ or $\mathbf{y} \subseteq \mathbf{z}$ where b_j, c_j are constant symbols and \mathbf{y}, \mathbf{z} are sequences of variables from $\{x_1, \dots, x_m\}$ with $|\mathbf{y}| = |\mathbf{z}| \leq k-1$.

- Assume first that ψ is of the form $x_i = b_j$ or $x_i = c_j$, and let $s \in Y^*$ be arbitrary. For (6), it suffices to show by Theorem 1 that

$$(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_s \psi. \tag{7}$$

First note that $s = h(f \circ t)$ for some automorphism $f \in \mathcal{F}$ and assignment $t \in Y$ for which, by the assumption and Theorem 1, $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_t \psi$. Hence for (7), we only need to show that $s(x_i) = t(x_i)$ in case $t(x_i)$ is listed in \mathbf{b} , and $s(x_i) = \varepsilon \circ t(x_i)$ in case $t(x_i)$ is listed in \mathbf{c} . For this, first recall that \mathcal{F} is the group generated by automorphisms $f_{\mathbf{a}}$ where $f_{\mathbf{a}}$ is obtained from item 4 of Theorem 8 and \mathbf{a} is a sequence listing $a_1, \dots, a_{k-1} \in A$ such that

$2 < \text{row}(a_i) < n - 1$, for $1 \leq i \leq k - 1$. Therefore f leaves all elementes in the first and the last row fixed when $f(\mathbf{b}) = \mathbf{b}$ and $f(\mathbf{c}) = \mathbf{c}$. On the other hand, by the definition of mid , $1 < \text{mid}(\text{row}(f \circ t(\mathbf{x}))) < n$, and hence $h(f \circ t)(x_i) = f \circ t(x_i)$ if $f \circ t(x_i)$ is in the first row, and $h(f \circ t)(x_i) = \varepsilon \circ f \circ t(x_i)$ if $f \circ t(x_i)$ is in the last row. Since tuples \mathbf{b} and \mathbf{c} are in the first and the last row, respectively, we conclude that the claim holds. The case where ψ is of the form $x_i = x_j$ or $\neg x_i = x_j$ is straightforward.

- Assume that ψ is of the form $E(x_i, x_j)$ or $\neg E(x_i, x_j)$. Again, let $s \in Y^*$ when $s = h(f \circ t)$ for some $f \in \mathcal{F}$ and $t \in Y$. For (7), consider first the case where

$$\text{row}(t(x_i)), \text{row}(t(x_j)) < \text{mid}(\text{row}(t(\mathbf{x}))), \text{ or} \quad (8)$$

$$\text{row}(t(x_i)), \text{row}(t(x_j)) > \text{mid}(\text{row}(t(\mathbf{x}))). \quad (9)$$

Since f is a row-preserving automorphism, we conclude by the definition of h that s maps both x_i and x_j either according to $f \circ t$ or according to $\varepsilon \circ f \circ t$. Since ε is also an automorphism, we obtain (7) in both cases. Assume then that (8) and (9) both fail. Then by symmetry suppose we have

$$\text{row}(t(x_i)) < \text{mid}(\text{row}(t(\mathbf{x}))) < \text{row}(t(x_j)).$$

Since $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_t \psi$, we have by item 1 of Theorem 8 that ψ is $\neg E(x_i, x_j)$. Since f and ε preserve the rows, we have

$$\text{row}(s(x_i)) < \text{mid}(\text{row}(s(\mathbf{x}))) < \text{row}(s(x_j)).$$

Therefore we obtain $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_s \neg E(x_i, x_j)$ which concludes this case.

- Assume that ϕ is $\mathbf{y} \subseteq \mathbf{z}$, for some $\mathbf{y} = y_1 \dots y_l$ and $\mathbf{z} = z_1 \dots z_l$ where $l \leq k - 1$. Let $s \in Y^*$ be arbitrary. For (6), we show that there exists a $s' \in Y^*$ such that $s(\mathbf{y}) = s'(\mathbf{z})$. Now $s = h(f \circ t)$ for some $f \in \mathcal{F}$ and $t \in Y$, and $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models_Y \psi$ by the assumption. Hence there exists a $t' \in Y$ such that $t(\mathbf{y}) = t'(\mathbf{z})$. Let now I list the indices $1 \leq i \leq l$ for which (i) or (ii) hold:⁵

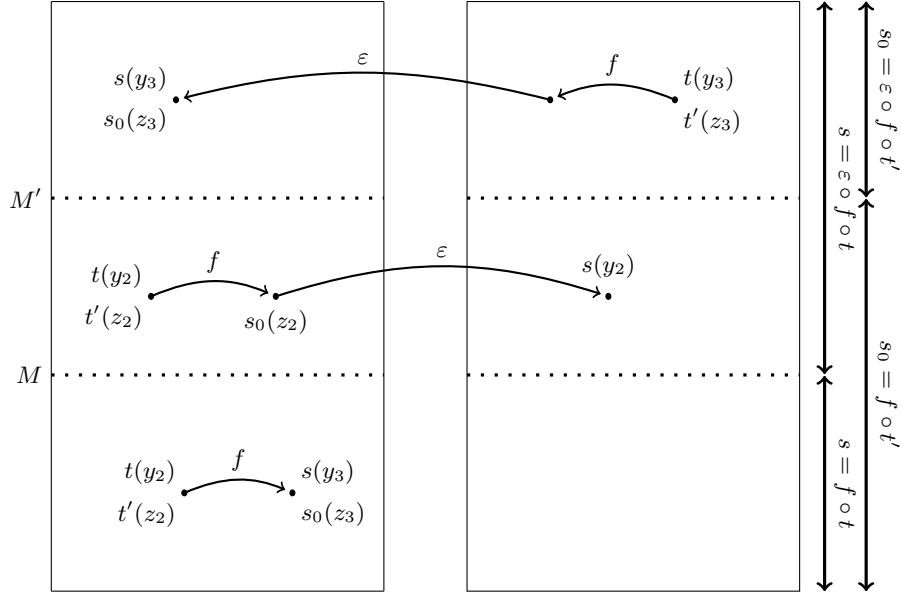
$$(i) \quad \text{row}(t(y_i)) < \text{mid}(\text{row}(t(\mathbf{x}))) \text{ and } \text{row}(t'(z_i)) > \text{mid}(\text{row}(t'(\mathbf{x}))),$$

$$(ii) \quad \text{row}(t(y_i)) > \text{mid}(\text{row}(t(\mathbf{x}))) \text{ and } \text{row}(t'(z_i)) < \text{mid}(\text{row}(t'(\mathbf{x}))).$$

Since $|I| \leq k - 1$, choosing $\mathbf{a} := (f \circ t'(z_i))_{i \in I}$ we find by item 4 of Theorem 8 an automorphism $f_{\mathbf{a}}$ that swaps $f \circ t'(z_i)$ to $\varepsilon \circ f \circ t'(z_i)$, for each $i \in I$, but leaves all elementes in rows of distance > 1 from $(\text{row}(f \circ t'(z_i)))_{i \in I}$ fixed. We now let $s' := h(f_{\mathbf{a}} \circ f \circ t')$. Since

$$1 < \text{mid}(\text{row}(t(\mathbf{x}))), \text{mid}(\text{row}(t'(\mathbf{x}))) < n$$

⁵ An example where $\mathbf{y} := y_1 y_2 y_3$ and $\mathbf{z} := z_1 z_2 z_3$ is illustrated in Fig. 2. Note that in the example, $I = \{2\}$ since the index number 2 satisfies (ii). Then letting $s_0 := h(f \circ t)$, we obtain $s(y_1 y_3) = s_0(z_1 z_3)$ but only $s(y_2) = \varepsilon \circ s_0(z_2)$. Fig. 3 shows that choosing $s' := h(f_{\mathbf{a}} \circ f \circ t')$, for $\mathbf{a} := f \circ t'(z_2)$, we obtain $s(\mathbf{y}) = s'(\mathbf{z})$.



$$\begin{cases} M' := \text{mid}(\text{row}(t'(\mathbf{x}))) \\ M := \text{mid}(\text{row}(t(\mathbf{x}))) \end{cases} \quad \begin{cases} s = h(f \circ t) \\ s_0 := h(f \circ t') \end{cases}$$

Fig. 2

by the definition, we have $2 < \text{row}(t'(z_i)) < n - 1$, for $i \in I$. Hence $f_{\mathbf{a}} \in \mathcal{F}$ and $s' \in Y^*$. Moreover, for $i \in I$, we obtain that

- (i) $s(y_i) = f \circ t(y_i) = f \circ t'(z_i) = \varepsilon \circ f_{\mathbf{a}} \circ f \circ t'(z_i) = s'(z_i)$, or
- (ii) $s(y_i) = \varepsilon \circ f \circ t(y_i) = \varepsilon \circ f \circ t'(z_i) = f_{\mathbf{a}} \circ f \circ t'(z_i) = s'(z_i)$.

For the first and last equalities note that $f_{\mathbf{a}}$ and f preserve the rows. For (i) recall also that ε is self-inverse.

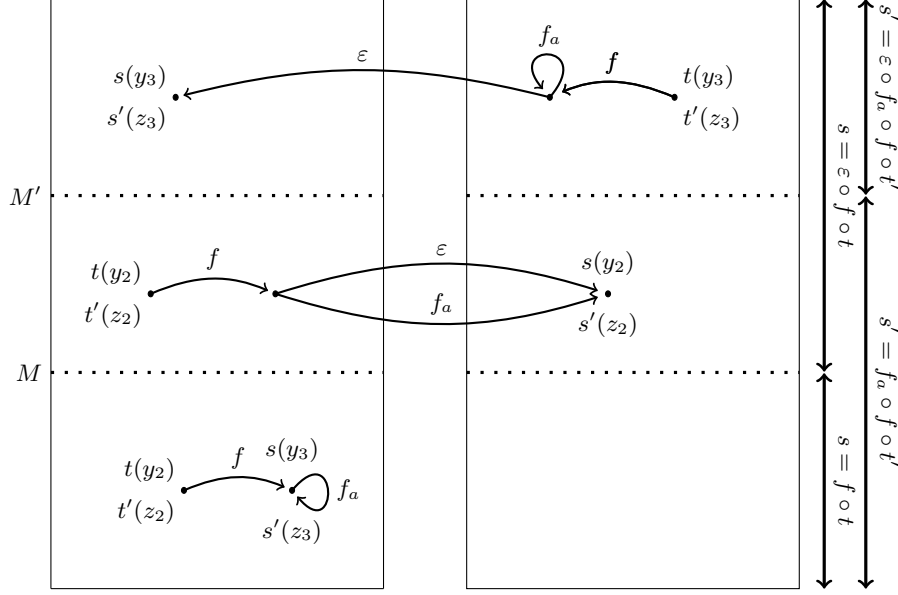
Let then $1 \leq j \leq l$ be such that $j \notin I$ when both (i) and (ii) and fail for j . Then we obtain

$$\text{row}(t(y_j)) > \text{mid}(\text{row}(t(\mathbf{x}))) \text{ and } \text{row}(t'(z_j)) > \text{mid}(\text{row}(t'(\mathbf{x}))), \text{ or} \quad (10)$$

$$\text{row}(t(y_j)) < \text{mid}(\text{row}(t(\mathbf{x}))) \text{ and } \text{row}(t'(z_j)) < \text{mid}(\text{row}(t'(\mathbf{x}))). \quad (11)$$

Assume first that (10) holds and let $i \in I$. Then either

- (i) $\text{row}(t(y_i)) < \text{mid}(\text{row}(t(\mathbf{x}))) < \text{row}(t(y_j))$, or
- (ii) $\text{row}(t'(z_i)) < \text{mid}(\text{row}(t'(\mathbf{x}))) < \text{row}(t'(z_j))$.



$$\begin{cases} M' := \text{mid}(\text{row}(t'(\mathbf{x}))) \\ M := \text{mid}(\text{row}(t(\mathbf{x}))) \end{cases} \quad \begin{cases} s = h(f \circ t) \\ s' := h(f_a \circ f \circ t'), \text{ for } a := f \circ t'(z_2) \end{cases}$$

Fig. 3

Since $t(y_j) = t'(z_j)$, $t(y_i) = t'(z_i)$, and f preserves the rows, in both cases we conclude that

$$|\text{row}(f \circ t'(z_j)) - \text{row}(f \circ t'(z_i))| > 1.$$

Therefore f_a leaves $f \circ t'(z_j)$ fixed. By (10) we now have

$$s(y_j) = \varepsilon \circ f \circ t(y_j) = \varepsilon \circ f \circ t'(z_j) = \varepsilon \circ f_a \circ f \circ t'(z_j) = s'(z_j).$$

The case where (11) holds is analogous. Hence $s(\mathbf{y}) = s'(\mathbf{z})$. This concludes the case of inclusion atom and thus the proof of $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{X^*} \theta$.

We have now concluded item 4 of the outline of the proof. Next we will show the last part of the proof. That is, we will show that X^* can be constructed by quantifying $Q^1 x_1 \dots Q^m x_m$ in \mathfrak{A} over $\{\emptyset\}$. For this, it suffices to show the following claim.

Claim. Let $a \in A$, $p \in \{1, \dots, m\}$ be such that $Q^p = \forall$, and $s \in X^* \upharpoonright \{x_1, \dots, x_{p-1}\}$. Then $s(a/x_p) \in X^* \upharpoonright \{x_1, \dots, x_p\}$.

Proof (Claim). Let a , p and s be as in the assumption. Then

$$s = h(f \circ t) \upharpoonright \{x_1, \dots, x_{p-1}\},$$

for some $f \in \mathcal{F}$ and $t \in X$. Let $a_0 = f^{-1}(a)$ and $a_1 = f^{-1} \circ \varepsilon(a)$. Note that both $t(a_0/x_p) \upharpoonright \{x_1, \dots, x_p\}$ and $t(a_1/x_p) \upharpoonright \{x_1, \dots, x_p\}$ are in $X \upharpoonright \{x_1, \dots, x_p\}$ since $Q^p = \forall$. Let $t_0, t_1 \in X$ extend $t(a_0/x_p) \upharpoonright \{x_1, \dots, x_p\}$ and $t(a_1/x_p) \upharpoonright \{x_1, \dots, x_p\}$, respectively. It suffices to show that either $h(f \circ t_0)$ or $h(f \circ t_1)$ (which both are in X^*) extend $s(a/x_p)$.

First note that since

$$t_0 \upharpoonright \{x_1, \dots, x_{p-1}\} = t_1 \upharpoonright \{x_1, \dots, x_{p-1}\} = t \upharpoonright \{x_1, \dots, x_{p-1}\}$$

we have by item 3 of the definition of mid that, for $i \leq p-1$, inequalities (12), (13) and (14) are equivalent:

$$\text{row}(t_0(x_i)) < \text{mid}(\text{row}(t_0(\mathbf{x}))), \quad (12)$$

$$\text{row}(t_1(x_i)) < \text{mid}(\text{row}(t_1(\mathbf{x}))), \quad (13)$$

$$\text{row}(t(x_i)) < \text{mid}(\text{row}(t(\mathbf{x}))). \quad (14)$$

Since also f preserves the rows, we have by the definition of h that $h(f \circ t_0)$, $h(f \circ t_1)$ and $h(f \circ t)$ all agree in variables x_1, \dots, x_{p-1} . Note that also ε preserves the rows, so have $\text{row}(a_0) = \text{row}(a_1)$. Since then $\text{row}(t_0(x_i)) = \text{row}(t_1(x_i))$, for $i \leq p$, we have by item 3 of the definition of mid that

$$\text{row}(t_0(x_p)) < \text{mid}(\text{row}(t_0(\mathbf{x}))) \text{ iff } \text{row}(t_1(x_p)) < \text{mid}(\text{row}(t_1(\mathbf{x}))).$$

Therefore, either

$$\text{row}(t_0(x_p)) < \text{mid}(\text{row}(t_0(\mathbf{x}))) \text{ or } \text{row}(t_1(x_p)) > \text{mid}(\text{row}(t_1(\mathbf{x}))).$$

Then in the first case $h(f \circ t_0)(x_p) = f \circ t_0(x_p) = a$, and in the second case $h(f \circ t_1)(x_p) = \varepsilon \circ f \circ t_1(x_p) = \varepsilon \circ \varepsilon(a) = a$. Hence $s(a/x_p) \in X^* \upharpoonright \{x_1, \dots, x_p\}$. This concludes the proof of the claim. \blacksquare

We have now showed that X^* can be constructed by quantifying $Q^1 x_1 \dots Q^m x_m$ in \mathfrak{A} over $\{\emptyset\}$. Also previously we showed that $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models_{X^*} \theta$. Therefore, since $\phi = Q^1 x_1 \dots Q^m x_m \theta$, we obtain that $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models \phi$. Hence the counter-assumption that $\phi(\mathbf{b}, \mathbf{c})$ defines $\neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ is false. Otherwise $\mathfrak{A} \models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \mathbf{c})$ would yield $(\mathfrak{A}, \mathbf{b}, \mathbf{c}) \models \phi$ from which $(\mathfrak{A}, \mathbf{b}, \varepsilon(\mathbf{c})) \models \phi$ follows. Therefore we would obtain $\mathfrak{A} \models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \varepsilon(\mathbf{c}))$ which contradicts with the fact that $\mathfrak{A} \not\models \neg[TC_{\mathbf{x}, \mathbf{y}} \text{EDGE}_k](\mathbf{b}, \varepsilon(\mathbf{c}))$ by Theorem 8. This concludes the proof of Lemma 1. \square

Theorem 6 follows now from Theorem 7 and Lemma 1.

4 Conclusion

We have showed that the arity fragments of inclusion logic give rise to an infinite expressivity hierarchy. Earlier, analogous results have been proved for dependence logic and independence logic. We also observed that the $\text{FO}(\subseteq)(k\forall)$ -hierarchy collapses at a very low level as it is the case with the $\text{FO}(\perp_c)(k\forall)$ -hierarchy. However, the $\text{FO}(=\dots)(k\forall)$ -hierarchy is strict since it can be related

to the strict $\text{ESO}_f(k\forall)$ -hierarchy. From the results of [3], [5] and this article, we obtain the following classification for syntactical hierarchies of dependence, independence and inclusion logic under the lax semantics.

	Arity of Dependency Atom	Number of \forall
$\text{FO}(=\dots)$	strict $\text{FO}(=\dots)(k\text{-dep}) <$ $\text{FO}(=\dots)(k+1\text{-dep})$	strict $\text{FO}(=\dots)(k\forall) <$ $\text{FO}(=\dots)(2k+2\forall)$
$\text{FO}(\perp_c)$	strict $\text{FO}(\perp_c)(k\text{-ind}) <$ $\text{FO}(\perp_c)(k+1\text{-ind})$	collapse at 2 $\text{FO}(\perp_c)(2\forall) = \text{FO}(\perp_c)$
$\text{FO}(\subseteq)$	strict $\text{FO}(\subseteq)(k\text{-inc}) <$ $\text{FO}(\subseteq)(k+1\text{-inc})$	collapse at 1 $\text{FO}(\subseteq)(1\forall) = \text{FO}(\subseteq)$

Since $\text{FO}(\subseteq)$ captures PTIME in finite ordered models, it would be interesting to investigate syntactical fragments of inclusion logic in that setting. It appears that then the techniques used in this article would be of no use. Namely, we cannot hope to construct two ordered models in the style of Theorem 8. In fixed point logics, this same question has been studied in the 90s. Imhof showed in [14] that the arity hierarchy of PFP remains strict in ordered models ($\text{PFP}^k <_{\mathcal{O}} \text{PFP}^{k+1}$) by relating the PFP^k -fragments to the degree hierarchy within PSPACE. For LFP and IFP, the same question appears to be more difficult, since both strictness and collapse have strong complexity theoretical consequences.

Theorem 9 ([14]). *For both IFP and LFP, collapse of arity hierarchy in ordered models implies $\text{PTIME} < \text{PSPACE}$, strictness implies $\text{LOGSPACE} < \text{PTIME}$.*

Proof. Sketch. For IFP, in case of collapse at k , the following chain of (in)equalities can be proved.

$$\text{PTIME} =_{\mathcal{O}} \text{IFP}^k \leq \text{PFP}^k <_{\mathcal{O}} \text{PFP} =_{\mathcal{O}} \text{PSPACE}.$$

For IFP, strictness implies

$$\text{LOGSPACE} =_{\mathcal{O}} \text{DTC} \leq_{\mathcal{O}} \text{IFP}^1 <_{\mathcal{O}} \text{IFP} =_{\mathcal{O}} \text{PTIME}.$$

For LFP the claim now follows from $\text{LFP}^k \leq \text{IFP}^k$ and $\text{IFP}^k \leq_{\mathcal{O}} \text{LFP}^{2k}$.

It might be possible to prove similar results for inclusion logic by relating the fragments $\text{FO}(\subseteq)(k\text{-inc})$ to arity fragments of fixed point logics. However, the translations between $\text{FO}(\subseteq)$ and GFP provided in [6] do not respect arities. It remains open whether collapse or strictness of the $\text{FO}(\subseteq)(k\text{-inc})$ -hierarchy have such strong consequences or whether it is possible to relate the $\text{FO}(\subseteq)(k\text{-inc})$ -fragments in ordered models to the degree hierarchy within PTIME? Another line would be to find some other syntactical parameter that would fit for this purpose.

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. Arnaud Durand and Juha Kontinen. Hierarchies in dependence logic. *ACM Transactions on Computational Logic (TOCL)*, 13(4):31, 2012.
4. 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.
5. Pietro Galliani, Miika Hannula, and Juha Kontinen. Hierarchies in independence logic. In Simona Ronchi Della Rocca, editor, *Computer Science Logic 2013 (CSL 2013)*, volume 23 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 263–280, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
6. Pietro Galliani and Lauri Hella. Inclusion Logic and Fixed Point Logic. In Simona Ronchi Della Rocca, editor, *Computer Science Logic 2013 (CSL 2013)*, volume 23 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 281–295, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
7. Pietro Galliani and Jouko A. Väänänen. On dependence logic. *CoRR*, abs/1305.5948, 2013.
8. Erich Grädel. Model-checking games for logics of imperfect information. *Theor. Comput. Sci.*, 493:2–14, 2013.
9. Erich Grädel and Jouko Väänänen. Dependence and independence. *Studia Logica*, 101(2):399–410, 2013.
10. Martin Grohe. Arity hierarchies. *Ann. Pure Appl. Logic*, 82(2):103–163, 1996.
11. Miika Hannula and Juha Kontinen. Hierarchies in independence and inclusion logic with strict semantics. *Manuscript*, 2014.
12. Wilfrid Hodges. Compositional Semantics for a Language of Imperfect Information. *Journal of the Interest Group in Pure and Applied Logics*, 5 (4):539–563, 1997.
13. Ehud Hrushovski. Extending partial isomorphisms of graphs. *Combinatorica*, 12(4):411–416, 1992.
14. Henrik Imhof. Computational aspects of arity hierarchies. In Dirk van Dalen and Marc Bezem, editors, *CSL*, volume 1258 of *Lecture Notes in Computer Science*, pages 211–225. Springer, 1996.
15. Jarmo Kontinen. Coherence and computational complexity of quantifier-free dependence logic formulas. In Juha Kontinen and Jouko Väänänen, editors, *Proceedings of Dependence and Independence in Logic*, pages 58–77. ESSLLI 2010, 2010.

16. Jouko Väänänen. *Dependence Logic*. Cambridge University Press, 2007.
17. Fan Yang. Expressing Second-order Sentences in Intuitionistic Dependence Logic. *Studia Logica*, 101(2):323–342, 2013.