Dependency Concepts up to Equivalence

Erich Grädel[1]
Mathematical Foundations of Computer Science, RWTH Aachen University, Aachen, Germany
graedel@logic.rwth-aachen.de

Matthias Hoelzel[2]
Mathematical Foundations of Computer Science, RWTH Aachen University, Aachen, Germany
hoelzel@logic.rwth-aachen.de

Abstract

Modern logics of dependence and independence are based on different variants of atomic dependency statements (such as dependence, exclusion, inclusion, or independence) and on team semantics: A formula is evaluated not with a single assignment of values to the free variables, but with a set of such assignments, called a team.

In this paper we explore logics of dependence and independence where the atomic dependency statements cannot distinguish elements up to equality, but only up to a given equivalence relation (which may model observational indistinguishabilities, for instance between states of a computational process or between values obtained in an experiment).

Our main goal is to analyse the power of such logics, by identifying equally expressive fragments of existential second-order logic or greatest fixed-point logic, with relations that are closed under the given equivalence. Using an adaptation of the Ehrenfeucht-Fraïssé method we further study conditions on the given equivalences under which these logics collapse to first-order logic, are equivalent to full existential second-order logic, or are strictly between first-order and existential second-order logic.

2012 ACM Subject Classification Theory of computation → Logic

Keywords and phrases Logics of dependence and independence, Team semantics, Existential second-order logic, Observational equivalence, Expressive power

Digital Object Identifier 10.4230/LIPIcs.CSL.2018.25

1 Introduction

Logics of dependence and independence (sometimes called logics of imperfect information) originally go back to the work of Henkin [9], Enderton [2], Walkoe [16], Blass and Gurevich [1], and others on Henkin quantifiers, whose semantics can be naturally described in terms of games of imperfect information. A next step in this direction have been the independence-friendly (IF) logics by Hintikka and Sandu [10] that incorporate explicit dependencies of quantifiers on each other and where again, the semantics is usually given in game-theoretic terms. For a detailed account on independence-friendly logics we refer to [13].

An important achievement towards the modern framework for logics of dependence and independence has been the model-theoretic semantics for IF-logics, due to Hodges [11], in terms of what he called *trumps*. This semantics is today called *team semantics*, where a *team* is understood as a set of assignments $s : \mathcal{V} \to A$, mapping a common finite domain of variables into the universe of a structure. The next step towards modern logics of

[1] This work has been initiated in a discussion between the first author and Jouko Väänänen during the Logical Structures in Computation Programme at the Simons Institute for Computing at UC Berkeley.
[2] Supported by DFG.

27th EACSL Annual Conference on Computer Science Logic (CSL 2018).
Editors: Dan Ghica and Achim Jung; Article No. 25; pp. 25:1–25:21
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

dependence and independence was the proposal by Väänänen [15] to consider dependencies as *atomic properties of teams* rather than stating them via annotations of quantifiers. He first introduced *dependence logic*, which is first-order logic on teams together with dependence atoms $\mathrm{dep}(x_1, \dots, x_m, y)$, saying that, in the given team, the variable y is functionally dependent on (i.e. completely determined by) the variables $x_1, \dots, x_m$. But there are many other atomic dependence properties that give rise to interesting logics based on team semantics. In [8] we have discussed the notion of independence (which is a much more delicate but also more powerful notion than dependence) and introduced *independence logics*, and Galliani [5] and Engström [3] have studied several logics with team properties based on notions originating in database dependency theory. The most important ones are *inclusion logic* $\mathrm{FO}(\subseteq)$, which extends first-order logic by atomic inclusion dependencies $(\bar{x} \subseteq \bar{y})$, which are true in a team X if every value for $\bar{x}$ in X also occurs as a value for $\bar{y}$ in X, and exclusion logic, based on exclusion statements $(\bar{x} \mid \bar{y})$, saying that $\bar{x}$ and $\bar{y}$ have disjoint sets of values in the team X. Exclusion logic has turned out to be equivalent to dependence logic [5].

Altogether this modern framework has lead to a genuinely new area in logic, with an interdisciplinary motivation of providing logical systems for reasoning about the fundamental notions of dependence and independence that permeate many scientific disciplines. Methods from several areas of computer science, including finite model theory, database theory, and the algorithmic analysis of games have turned out as highly relevant for this area. For more information, we refer to the volume [4] and the references there.

In this paper we explore logics that are based on weaker variants of dependencies. We consider atomic dependence statements that do not distinguish elements up to equality, but only up to coarser equivalencies. This is motivated by the quite familiar situation in many applications that elements, such as for instance states in a computation or values obtained in experiments, are subject to observational indistinguishabilities, which we model here via an equivalence relation $\approx$ on the set of possible values. For *dependence atoms* $\mathrm{dep}_{\approx}(\bar{x}, y)$ this means that we can say that whenever the values of $\bar{x}$ are indistinguishable for certain assignments in a team, then so are the values of y. Similarly an exclusion statement between x and y, up to an equivalence relation $\approx$, says that no value for x in the team is equivalent to a value of y, and an *inclusion statement* $x \subseteq_{\approx} y$ means that every value for x is equivalent to some value for y. Finally, the most powerful of such notions, independence of x and y up to equivalence, means that additional information about the equivalence class of the value of one variable does not help to learn anything new about the value of the other, or to put it differently, whenever a value a for x and a value b for y occur in the team, then there is an assignment in the team whose value for x is equivalent to a and whose value for y is equivalent to b.

Formal definitions of these dependencies, extended to tuples of variables, will be given in the next section.

The main goal of this paper is to *understand the expressive power of the logics with dependencies up to equivalence.* In general, logical operations on teams have a second-order nature, and indeed, dependencies and team semantics may take the power of first-order logic FO up to existential second-order logic Σ^1_1. To make this precise we recall the *standard translation*, due to [15, 12], from formulae with team semantics into sentences of existential second-order logic.

First of all, we have to keep in mind the different nature of team semantics and classical Tarski semantics. For a formula with team semantics, we write $\mathfrak{A} \models_X \varphi$ to denote that φ is true in the structure $\mathfrak{A}$ for the team X, and for classical Tarski semantics we write $\mathfrak{A} \models_s \varphi$ to denote that φ is true in $\mathfrak{A}$ for the assignment s.

For formulae with free variables the translation from a logic with team semantics into one with Tarski semantics requires that we represent the team in some way. The standard way to do this is by identifying a team X of assignments $s : \{x_1, \dots x_k\} \to A$ with the relation $\{s(x_1, \dots, x_k) \in A^k : s \in X\} \subseteq A^k$ which, by slight abuse of notation, we also denote by X. One then translates formulae $\varphi(x_1, \dots, x_k)$ of vocabulary τ of a logic L with team semantics into *sentences* φ^* in Σ^1_1 of the expanded vocabulary $\tau \cup \{X\}$ such that for every structure $\mathfrak{A}$ and every team X we have that

$$\mathfrak{A} \models_X \varphi(x_1, \dots, x_k) \iff (\mathfrak{A}, X) \models \varphi^*.$$

To illustrate this second-order nature we recall the meaning of disjunctions and existential quantifications in team semantics, and their standard translation into Σ^1_1. Disjunctions split the team, i.e..

$$\mathfrak{A} \models_X \psi \vee \varphi \quad :\iff \quad X = Y \cup Z \text{ such that } \mathfrak{A} \models_Y \psi \text{ and } \mathfrak{A} \models_Z \varphi$$

which leads to the translation $(\psi \vee \varphi)^*(X) := \exists Y \exists Z(X = Y \cup Z \wedge \psi^*(Y) \wedge \varphi^*(Z))$. Existential quantification requires the extension of the given team by providing for each of its assignments a *non-empty set of witnesses* for quantified variables, i.e.,

$$\mathfrak{A} \models_X \exists y \psi \quad :\iff \quad \text{there exists a function } F : X \to \mathcal{P}^+(A) \text{ such that } \mathfrak{A} \models_{X[y \mapsto F]} \psi$$

where $X[y \mapsto F]$ is the set of all assignments $s[y \mapsto a]$ that update an assignment $s \in X$ by mapping y to some value $a \in F(s)$. This leads to the translation $(\exists y \psi)^*(X) := \exists Y \forall \bar{x}((X\bar{x} \leftrightarrow \exists y Y \bar{x} y) \wedge \psi^*(Y))$.

Some remarks are in order: One may wonder why it is appropriate to provide a non-empty set of witnesses for an existentially quantified variable rather than just a single witness as in standard Tarski semantics for first-order logic. Indeed there are many cases where a single witness, i.e. a function $F : X \to A$ rather than $F : X \to \mathcal{P}^+(A)$ suffices, in fact in all cases where the logic is *downwards closed*, i.e. when $\mathfrak{A} \models_X \psi$ implies that also $\mathfrak{A} \models_Y \psi$ for all subteams $Y \subseteq X$. Examples of downwards closed logics are dependence logic and exclusion logic. However, for logics that are not downwards closed, such as inclusion logic and independence logic, the so-called strict semantics requiring single witnesses of existentially quantified variables leads to pathologies such as non-locality: the meaning of a formula might depend on the values of variables that do not even occur in it. A second relevant remark is that all the logics considered here have the *empty team property*: For all sentences φ and all structures $\mathfrak{A}$, we have that $\mathfrak{A} \models_\varnothing \varphi$. To evaluate sentences (formulae without free variables) we therefore have to consider not the empty team, but the team $\{\varnothing\}$ consisting just of the empty assignment. For a sentence ψ we write $\mathfrak{A} \models \psi$ if $\mathfrak{A} \models_{\{\varnothing\}} \psi$.

On the basis of the standard translation in Σ^1_1 we can say that we understand the expressive power of a first-order logic with dependencies, when we have identified the fragment $\mathcal{F}$ of existential second-order logic which is equivalent in the sense just described. The following is known in this context:

(1) Dependence logic and exclusion logic are equivalent to the fragment of all Σ^1_1-sentences $\psi(X)$ in which the predicate X describing the team appears only negatively [12].

(2) Independence logic and inclusion-exclusion logic are equivalent with full Σ^1_1 (and thus can describe all NP-properties of teams) [5].

(3) The extension of FO by inclusion and exclusion atoms of single variables only (not tuples of variables) is equivalent to monadic Σ^1_1 [14].

(4) First-order logic without any dependence atoms has the so-called *flatness property*: $\mathfrak{A} \models_X \varphi \iff \mathfrak{A} \models_s \varphi$ for all $s \in X$. It thus corresponds to a very small fragment of Σ^1_1, namely FO-sentences of form $\forall \bar{x}(X\bar{x} \to \varphi(\bar{x}))$ where $\varphi(\bar{x})$ does not contain X.

(5) Inclusion logic FO($\subseteq$) corresponds to GFP^+, the fragment of fixed-point logic that uses only (non-negated) greatest fixed-points. Since a greatest fixed-point formula $[\mathbf{gfp}\, R\bar{x}\,.\,\psi(R,\bar{x})](\bar{y})$ readily translates into $(\exists R)((\forall \bar{x}(R\bar{x} \to \psi(R,\bar{x})) \wedge R\bar{y}))$, GFP^+ can be viewed as a fragment of Σ^1_1. Galliani and Hella [6] established that inclusion logic is equivalent to the set of sentences of form $\forall \bar{x}(X\bar{x} \to \psi(X,\bar{x}))$, where $\psi(X,\bar{x})$ is a formula in GFP^+ in which X occurs only positively. A different proof for this result, based on safety games and game interpretations, has been presented in [7].

Hence the question arises how these fragments change when the standard dependency notions are replaced by dependencies up to equivalence. There is a natural conjecture: *One has to restrict existential second-order quantification to relations that are closed under the given equivalence relation, i.e. to relations that can be written as unions of equivalence classes* (where equivalence is extended to tuples component-wise). We denote the resulting variant of existential second-order logic by $\Sigma^1_1(\approx)$.

Notice however, that to decide this conjecture is far from being trivial and, in fact, *the restriction of the standard translation to quantification over $\approx$-closed relations fails.* Even for simple disjunctions, the existential second-order expression given above describing the split of the team will not work anymore once we restrict quantification to $\approx$-closed relations because we cannot assume that the relevant subteams are $\approx$-closed. Here is a simple example, not even involving any dependencies: Consider the formula $x = y \vee x \neq y$ which is trivially true in any team X, by the split $X = Y \cup Z$ where Y contains the assignments s which $s(x) = s(y)$ and $Z = X \setminus Y$ (and this is the only split that works). However if there are elements $a \neq b$ with $a \approx b$ then in general neither Y nor Z are $\approx$-closed, even if X is.

Nevertheless we shall prove that the conjecture is true, and that we can characterize the expressive power of dependence logics up to equivalence by appropriate fragments of $\Sigma^1_1(\approx)$. This is based on a much more sophisticated translation from logics with team semantics into existential second-order logic that adapts ideas from [14]. We shall also present a fragment of GFP^+ that has the same expressive power as inclusion logic up to equivalence.

Our next question is then how the expressive power of $\Sigma^1_1(\approx)$, and hence logics of dependence up to equivalence, compare to first-order logic and to full Σ^1_1. Of course this depends on the properties of the underlying equivalence relation, notably on the number and sizes of its equivalence classes.

(1) On any class of structures on which $\approx$ has only a bounded number of equivalence classes, $\Sigma^1_1(\approx)$, and hence all logics with dependencies up to equivalence as well, collapse to FO.

(2) On any class of structures in which all equivalence classes have bounded size, and only a bounded number of classes have more than one element, $\Sigma^1_1(\approx) \equiv \Sigma^1_1$.

(3) In general, and in particular on the classes of structures where all equivalence classes have size at most k (for $k > 1$), or that have only a bounded number of equivalence classes of size >1, the expressive power of $\Sigma^1_1(\approx)$, and all the considered logics of dependence up to equivalence, are strictly between FO and Σ^1_1.

To prove this we shall use appropriate variants of Ehrenfeucht-Fraïssé games for these logics.

2 The Logics $\mathrm{FO}(\Omega_\approx)$ and $\Sigma_1^1(\approx)$

Let τ be a signature containing a binary relation symbol $\approx$ and let $(\tau, \approx)$ denote the class of τ-structures $\mathfrak{A}$ in which $\approx$ is interpreted by an equivalence relation on the universe A of $\mathfrak{A}$. For every $\mathfrak{A} \in (\tau, \approx)$ and every $\bar{a}, \bar{b} \in A^n$ we write $\bar{a} \approx \bar{b}$, if $a_i \approx b_i$ for every $i \in \{1, \dots, n\}$. Given two relations $R, S \subseteq A^k$ of the same arity we write $R \subseteq_\approx S$ if for every $\bar{a} \in R$, there exists some $\bar{b} \in S$ with $\bar{a} \approx \bar{b}$. We further write $R \approx S$ if $R \subseteq_\approx S$ and $S \subseteq_\approx R$. Furthermore, we define the $\approx$-closure of R as $R_\approx := \{\bar{a} : \bar{a} \approx \bar{b} \text{ for some } \bar{b} \in R\}$ and say that R is $\approx$-closed if, and only if, $R = R_\approx$.

A team over $\mathfrak{A}$ is a set X of assignments $s : \mathrm{dom}(X) \to A$ mapping a common finite domain of variables into the universe A of $\mathfrak{A}$. Given a tuple $\bar{y}$ of variables from $\mathrm{dom}(X)$, we denote by $X(\bar{y}) := \{s(\bar{y}) : s \in X\}$ the set of values that $\bar{y}$ takes in X. The semantics of (in)dependence, inclusion and exclusion atoms up to $\approx$ is given as follows:

▶ **Definition 1.** Let X be a team over $\mathfrak{A}$. Then we define

$$\begin{aligned}
\mathfrak{A} \models_X \mathrm{dep}_\approx(\bar{x}, y) \quad &:\Longleftrightarrow \quad \text{for all } s, s' \in X, \text{ if } s(\bar{x}) \approx s'(\bar{x}) \text{ then also } s(y) \approx s'(y),\\
\mathfrak{A} \models_X \bar{x} \perp_\approx \bar{y} \quad &:\Longleftrightarrow \quad \text{for all } s, s' \in X \text{ there exists some } s'' \in X \text{ such that}\\
&\qquad\quad s''(\bar{x}) \approx s(\bar{x}) \text{ and } s''(\bar{y}) \approx s'(\bar{y}),\\
\mathfrak{A} \models_X \bar{x} \subseteq_\approx \bar{y} \quad &:\Longleftrightarrow \quad X(\bar{x}) := \{s(\bar{x}) : s \in X\} \subseteq_\approx X(\bar{y}),\\
\mathfrak{A} \models_X \bar{x} \mid_\approx \bar{y} \quad &:\Longleftrightarrow \quad s(\bar{x}) \not\approx s'(\bar{y}) \text{ for all } s, s' \in X.
\end{aligned}$$

For $\Omega_\approx \subseteq \{\mathrm{dep}_\approx, \perp_\approx, \subseteq_\approx, |_\approx\}$ we denote by $\mathrm{FO}(\Omega_\approx)$ the set of all first-order formulas in negation normal form where we additionally allow positive occurrences of $\Omega_\approx$-atoms. The semantics of first-order literals and of the logical operators are the usual ones in (lax) team semantics:

▶ **Definition 2.** Let $\varphi_1, \varphi_2, \psi \in \mathrm{FO}(\Omega_\approx)$, ϑ be some first-order literal and X a team over $\mathfrak{A}$.

$$\begin{aligned}
\mathfrak{A} \models_X \vartheta \quad &:\Longleftrightarrow \quad \mathfrak{A} \models_s \vartheta \text{ for every } s \in X,\\
\mathfrak{A} \models_X \varphi_1 \wedge \varphi_2 \quad &:\Longleftrightarrow \quad \mathfrak{A} \models_X \varphi_1 \text{ and } \mathfrak{A} \models_X \varphi_2\\
\mathfrak{A} \models_X \varphi_1 \vee \varphi_2 \quad &:\Longleftrightarrow \quad X \text{ can be represented as } X = X_1 \cup X_2 \text{ such that}\\
&\qquad\quad \mathfrak{A} \models_{X_1} \varphi_1 \text{ and } \mathfrak{A} \models_{X_2} \varphi_2\\
\mathfrak{A} \models_X \forall x \psi \quad &:\Longleftrightarrow \quad \mathfrak{A} \models_{X[x \mapsto A]} \psi\\
\mathfrak{A} \models_X \exists x \psi \quad &:\Longleftrightarrow \quad \mathfrak{A} \models_{X[x \mapsto F]} \psi \text{ for some } F : X \to \mathcal{P}(A) \setminus \{\varnothing\}.
\end{aligned}$$

Here we have $X[x \mapsto A] := \{s[x \mapsto a] : s \in X, a \in A\}$ and $X[x \mapsto F] := \{s[x \mapsto a] : s \in X, a \in F(s)\}$. Sometimes we shall call a team Y an $\{x\}$-extension of X, if $Y = X[x \mapsto F]$ for some function $F : X \to \mathcal{P}(A) \setminus \{\varnothing\}$.

Many standard results concerning the closure properties and relationships between different logics of dependence and independence (see e.g. [5]) carry over to this new setting with equivalences, by easy and straightforward adaptations of proofs (which are therefore omitted here). In particular, this includes the following observations:

- For all formulae in these logics the *locality principle* holds: $\mathfrak{A} \models_X \varphi$ if, and only if, $\mathfrak{A} \models_{X \restriction \mathrm{free}(\varphi)} \varphi$ (where $X \restriction \mathrm{free}(\varphi) := \{s \restriction \mathrm{free}(\varphi) : s \in X\}$ is the restriction of X to the free variables of φ).
- The logics $\mathrm{FO}(\mathrm{dep}_\approx)$ and $\mathrm{FO}(|_\approx)$ are equivalent and downwards closed.
- The logic $\mathrm{FO}(\subseteq_\approx)$ is closed under unions of teams, and incomparable with $\mathrm{FO}(\mathrm{dep}_\approx)$ and $\mathrm{FO}(|_\approx)$.
- Independence logic with equivalences, $\mathrm{FO}(\perp_\approx)$, is equivalent to inclusion-exclusion logic with equivalences, $\mathrm{FO}(\subseteq_\approx, |_\approx)$.

A much more difficult problem is to understand the expressive power of these logics in connection with existential second-order logic Σ^1_1. As mentioned above, formulae of independence logic or, equivalently, inclusion-exclusion logic (without equivalences) have the same expressive power as existential second-order sentences, and weaker logics such as dependence logic, exclusion logic, or inclusion logic correspond to fragments of Σ^1_1. To describe the expressive power of dependence logics with equivalences we introduce the $\approx$-closed fragment $\Sigma^1_1(\approx)$ of Σ^1_1 and show that it captures the expressiveness of $\mathrm{FO}(\subseteq_\approx, |_\approx)$.

▶ **Definition 3.** The logic $\Sigma^1_1(\approx)$ consists of sentences of the form

$$\psi := \exists_\approx R_1 \dots \exists_\approx R_k \varphi(R_1, \dots, R_k)$$

where $\varphi \in \mathrm{FO}(\tau \cup \{R_1, \dots, R_k\})$. The semantics of ψ is given in terms of $\approx$-closed relations:

$$\mathfrak{A} \models \psi :\Longleftrightarrow \text{there are } \approx\text{-closed relations } R_1, \dots, R_k \text{ such that } (\mathfrak{A}, R_1, \dots, R_k) \models \varphi.$$

3 The Expressive Power of $\mathrm{FO}(\subseteq_\approx, |_\approx)$

In this section we establish that $\mathrm{FO}(\subseteq_\approx, |_\approx)$ has exactly the expressive power of $\Sigma^1_1(\approx)$. This means that every formula $\varphi(\overline{x}) \in \mathrm{FO}(\subseteq_\approx, |_\approx)$ can be translated into an equivalent sentence $\varphi' \in \Sigma^1_1(\approx)$ using an additional predicate for the team such that

$$\mathfrak{A} \models_X \varphi(\overline{x}) \Longleftrightarrow (\mathfrak{A}, X) \models \varphi'(X).$$

Conversely, we are also going to show how a given sentence $\psi \in \Sigma^1_1(\approx)$ can be translated into an equivalent sentence $\psi^+ \in \mathrm{FO}(\subseteq_\approx, |_\approx)$.

3.1 From $\Sigma^1_1(\approx)$ to $\mathrm{FO}(\subseteq_\approx, |_\approx)$

To capture the semantics of a sentence $\exists_\approx R_1 \dots \exists_\approx R_k \varphi \in \Sigma^1_1(\approx)$ in $\mathrm{FO}(\subseteq_\approx, |_\approx)$ we adapt ideas by Rönnholm [14] and use tuples of variables $\overline{v_1}, \dots, \overline{v_k}$ of length $|\overline{v_i}| = \mathrm{ar}(R_i)$ in order to simulate the ($\approx$-closed) relations $R_1, \dots, R_k$. The reason why this is possible lies in the fact that we are using team semantics: In a given team X with $\{\overline{v_1}, \dots, \overline{v_k}\} \subseteq \mathrm{dom}(X)$ we naturally have that $X(\overline{v_i})$ corresponds to a (not necessarily $\approx$-closed) relation. The most important step is to find a formula $\varphi^\star(\overline{v_1}, \dots, \overline{v_k}) \in \mathrm{FO}(\subseteq_\approx, |_\approx)$ such that

$$(\mathfrak{A}, \overline{R}) \models \varphi \Longleftrightarrow \mathfrak{A} \models_X \varphi^\star(\overline{v_1}, \dots, \overline{v_k})$$

where $X = \{s : s(\overline{v_i}) \in R_i \text{ for every } i \in \{1, \dots, k\}\}$. Towards this end, $\varphi^\star$ is constructed (inductively) while using inclusion/exclusion atoms to express (non)membership in $R_1, \dots, R_k$. For example, $\overline{x} \subseteq_\approx \overline{v_i}$ means that $s(\overline{x}) \in X(\overline{v_i})_\approx = R_i$ for every $s \in X$, while $\overline{x} \mid_\approx \overline{v_i}$ expresses that $s(\overline{x}) \notin X(\overline{v_i})_\approx = R_i$ for every $s \in X$. Therefore, the semantics of $R_i\overline{x}$ resp. $\neg R_i \overline{x}$ is captured by $\overline{x} \subseteq_\approx \overline{v_i}$ resp. $\overline{x} \mid_\approx \overline{v_i}$. But of course, it could be the case that φ is a much more complicated formula made up of quantifiers, conjunction or disjunctions. It turns out that quantifiers and conjunction can be handled with ease by simply setting

$$(Qu\vartheta)^\star := Qu(\vartheta^\star) \text{ for both quantifiers } Q \in \{\exists, \forall\}, \text{ and}$$
$$(\vartheta_1 \wedge \vartheta_2)^\star := \vartheta_1^\star \wedge \vartheta_2^\star,$$

because when evaluating conjunctions in team semantics, the team is not modified and in the process of evaluating quantifiers there are just more columns added to the team (w.l.o.g. we assume that every variable in the formula occurs either freely or is quantified exactly once). However, for disjunctions the situation is much more delicate because it is *not* possible to define $(\vartheta_1 \vee \vartheta_2)^\star$ as $\vartheta_1^\star \vee \vartheta_2^\star$. The reason for this is that after splitting the team X into X_1, X_2 with $X = X_1 \cup X_2$ and $\mathfrak{A} \models_{X_j} \vartheta_j^\star$ it cannot be guaranteed that $X_j(\overline{v_i})$ still describes the original R_i (up to equivalence). To make sure that we do not loose information about $R_1, \ldots, R_k$, we use instead an adaptation of the *value preserving disjunction* that was introduced by Rönnholm [14].

▶ **Lemma 4.** *Let $\psi_1, \psi_2 \in \mathrm{FO}(\subseteq_\approx, |_\approx)$ and $\overline{v_1}, \ldots, \overline{v_k}$ be some tuples of variables. Then there exists a formula $\psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee} \psi_2 \in \mathrm{FO}(\subseteq_\approx, |_\approx)$ such that the following are equivalent:*

(i) $\mathfrak{A} \models_X \psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee} \psi_2$

(ii) *$X = X_1 \cup X_2$ for some teams X_1, X_2 such that for both $j = 1$ and $j = 2$:*
- *$\mathfrak{A} \models_{X_j} \psi_j$, and*
- *if $X_j \neq \varnothing$, then $X_j(\overline{v_i}) \approx X(\overline{v_i})$ for all $i \in \{1, \ldots, k\}$.*

Proof. The construction of $\psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee} \psi_2$ relies on the *intuitionistic disjunction* $\psi_1 \sqcup \psi_2$ with

$$\mathfrak{A} \models_X \psi_1 \sqcup \psi_2 \iff \mathfrak{A} \models_X \psi_1 \text{ or } \mathfrak{A} \models_X \psi_2.$$

On structures $\mathfrak{A} \in (\tau, \approx)$ with $\approx^{\mathfrak{A}} \neq A^2$ this is definable in $\mathrm{FO}(\subseteq_\approx, |_\approx)$ since

$$\psi_1 \sqcup \psi_2 \equiv \exists c_\ell \exists c_r (\mathrm{dep}_\approx(c_\ell) \wedge \mathrm{dep}_\approx(c_r) \wedge [(c_\ell \approx c_r \wedge \psi_1) \vee (\neg c_\ell \approx c_r \wedge \psi_2)])$$

where c_ℓ and c_r are some variables not occurring in ψ_1 or ψ_2. Note that $\mathrm{dep}_\approx(c)$ expresses that c only assumes values from a single equivalence class. The proof of this equivalence is a simple exercise. Now consider the following formula, which is a modification of a construction by Rönnholm [14].

$$\begin{aligned}\psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee'} \psi_2 \coloneqq (\psi_1 \sqcup \psi_2) \sqcup \exists c_\ell \exists c_r \Big(&\mathrm{dep}_\approx(c_\ell) \wedge \mathrm{dep}_\approx(c_r) \wedge c_\ell \not\approx c_r \wedge \\ &\exists y \Big([(y \approx c_\ell \wedge \psi_1) \vee (y \approx c_r \wedge \psi_2)] \wedge \\ &\bigwedge_{i=1}^{k} \Theta_i \wedge \Theta_i' \Big)\Big).\end{aligned}$$

Θ_i and Θ_i' are given by

$$\begin{aligned}\Theta_i &\coloneqq \exists \overline{z_1} \exists \overline{z_2} ([(y \approx c_\ell \wedge \overline{z_1} = \overline{v_i} \wedge \overline{z_2} = \overline{c_\ell}) \vee (y \approx c_r \wedge \overline{z_1} = \overline{c_\ell} \wedge \overline{z_2} = \overline{v_i})] \\ &\qquad \wedge \overline{v_i} \subseteq_\approx \overline{z_1} \wedge \overline{v_i} \subseteq_\approx \overline{z_2}) \\ \Theta_i' &\coloneqq \exists \overline{z_1} \exists \overline{z_2} ([(y \approx c_\ell \wedge \overline{z_1} = \overline{v_i} \wedge \overline{z_2} = \overline{c_r}) \vee (y \approx c_r \wedge \overline{z_1} = \overline{c_r} \wedge \overline{z_2} = \overline{v_i})] \\ &\qquad \wedge \overline{v_i} \subseteq_\approx \overline{z_1} \wedge \overline{v_i} \subseteq_\approx \overline{z_2}).\end{aligned}$$

where $\overline{c_\ell} = (c_\ell, c_\ell, \ldots, c_\ell)$ and $\overline{c_r} = (c_r, c_r, \ldots, c_r)$ are always tuples of the correct length.

It is not difficult to prove that this formula is almost what we want: it satisfies the properties required by Lemma 4 under the additional condition that $\approx$ has at least two different equivalence classes. To get rid of this condition, we put:

$$\psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee} \psi_2 \coloneqq [\forall x \forall y (x \approx y) \wedge (\psi_1 \vee \psi_2)] \vee \left[\exists x \exists y (x \not\approx y) \wedge \left(\psi_1 \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee'} \psi_2\right)\right]. \quad ◀$$

We can now complete the inductive definition of $\varphi^\star$ by:

$$(\vartheta_1 \vee \vartheta_2)^\star \coloneqq \vartheta_1^\star \underset{\overline{v_1},\ldots,\overline{v_k}}{\vee} \vartheta_2^\star$$

By rather straighforward inductions one can establish the following two lemmata.

▶ **Lemma 5.** *For every $\mathfrak{A} \in (\tau, \approx)$ and every team X with $dom(X) = \{\overline{v_1}, \ldots, \overline{v_k}\}$,*

$$\mathfrak{A} \models_X \varphi^\star \Longrightarrow (\mathfrak{A}, R_1^X, \ldots, R_k^X) \models_X \varphi$$

where $R_i^X := (X(\overline{v_i}))_\approx$ for $i = 1, \ldots, k$, i.e. R_i^X is defined as the $\approx$-closure of $X(\overline{v_i})$.

▶ **Lemma 6.** *Let $\overline{R} = (R_1, \ldots, R_k)$ be a tuple of non-empty $\approx$-closed relations with $(\mathfrak{A}, \overline{R}) \models \varphi$. Then $\mathfrak{A} \models_Y \varphi^\star$ where $Y := \{(\overline{v_1}, \ldots, \overline{v_k}) \mapsto (\overline{a_1}, \ldots, \overline{a_k}) : \overline{a_1} \in R_1, \ldots, \overline{a_k} \in R_k\}$.*

The non-emptiness requirement of $R_1, \ldots, R_k$ does not create a serious problem, because by rewriting the formula φ it can be assumed w.l.o.g. that $\exists_\approx R_1 \ldots \exists_\approx R_k \varphi$ is satisfied in a structure $\mathfrak{A}$ if, and only if, there are *non-empty* $\approx$-closed relations $R_1, \ldots, R_k$ such that $(\mathfrak{A}, \overline{R}) \models \varphi$.

▶ **Theorem 7.** *$\psi := \exists \overline{v_1} \ldots \exists \overline{v_k} \varphi^\star$ is equivalent to $\exists_\approx R_1 \ldots \exists_\approx R_k \varphi$.*

Proof. Assume that $\mathfrak{A} \models \psi$. It follows that there exists some team X with $\mathrm{dom}(X) = \{\overline{v_1}, \ldots, \overline{v_k}\}$ and $\mathfrak{A} \models_X \varphi^\star$. By Lemma 5 and $\mathrm{free}(\varphi) = \varnothing$, we obtain that $(\mathfrak{A}, R_1^X, \ldots, R_k^X) \models \varphi$. By definition, the relations R_i^X are $\approx$-closed and, hence, $\mathfrak{A} \models \exists_\approx R_1 \ldots \exists_\approx R_k \varphi$.

For the converse direction, let $\mathfrak{A} \models \exists_\approx R_1 \ldots \exists_\approx R_k \varphi$. So there exists some *non-empty* $\approx$-closed relations $R_1, \ldots, R_k$ such that $(\mathfrak{A}, R_1, \ldots, R_k) \models \varphi$. So by Lemma 6, it follows that $\mathfrak{A} \models_Y \varphi^\star$ where Y is the team given in Lemma 6. This leads to $\mathfrak{A} \models \exists \overline{v_1} \ldots \exists \overline{v_k} \varphi^\star$. ◀

3.2 From $\mathrm{FO}(\subseteq_\approx, |_\approx)$ to $\Sigma_1^1(\approx)$

Up to this point we only know that $\Sigma_1^1(\approx) \leq \mathrm{FO}(\subseteq_\approx, |_\approx)$. In this section we prove that these two logics have in fact the same expressive power. Towards this end, we demonstrate how a given formula $\varphi \in \mathrm{FO}(\subseteq_\approx, |_\approx)$ can be translated into $\Sigma_1^1(\approx)$. There are two obstacles that we need to overcome:

1. When viewed as relations, teams usually are not $\approx$-closed, so we cannot use the quantifier $\exists_\approx$ to fetch the subteams we would need to satisfy the subformulae of e.g. a disjunction.
2. Unlike in Σ_1^1, where a formula of the form $\forall x \exists Y (\ldots)$ is equivalent to formula like $\exists Y' \forall x (\ldots)$ where $\mathrm{ar}(Y') = \mathrm{ar}(Y) + 1$, there seems to be no obvious way to perform a similar syntactic manipulation in $\Sigma_1^1(\approx)$. Thus we have to be content with the limited quantification that $\Sigma_1^1(\approx)$ allows us.

The main idea of the construction, which is inspired by [14], is to replace every inclusion and exclusion atom ϑ by a seperate new relation symbol R_ϑ that contains certain values enabling us to express the semantics of φ in $\Sigma_1^1(\approx)$.

First we describe how this approach deals with exclusion atoms. Let $\vartheta_1, \ldots, \vartheta_k$ be an enumeration of all occurrences of exclusion atoms $\vartheta_i = \overline{u_i} \mid_\approx \overline{w_i}$ in φ. We assume w.l.o.g. that the tuples $\overline{u_1}, \ldots, \overline{u_k}, \overline{w_1}, \ldots, \overline{w_k}$ are pairwise different. We use new relation symbols $R_{\vartheta_1}, \ldots, R_{\vartheta_k}$ that are intended to separate the sets of possible values for $\overline{v_i}$ and $\overline{w_i}$ (up to equivalence). The desired translation $\varphi^\star$ of φ is now obtained by replacing the exclusion atoms $\vartheta_i = \overline{u_i} \mid_\approx \overline{w_i}$ by $R_{\vartheta_i} \overline{u_i} \wedge \neg R_{\vartheta_i} \overline{w_i}$. This construction leads to the following result. The proof is by induction over φ and is given in the appendix.

▶ **Theorem 8.** *For every formula $\varphi(\overline{x}) \in \mathrm{FO}(|_\approx, \subseteq_\approx)$ with signature τ there exists a formula $\varphi^\star(\overline{x}) \in \mathrm{FO}(\subseteq_\approx)$ with signature $\tau \cup \{\overline{R}\}$, where $\overline{R}$ is a tuple of new relation symbols, such that for every $\mathfrak{A} \in (\tau, \approx)$ and every team X the following are equivalent:*

(i) $\mathfrak{A} \models_X \varphi$

(ii) *There are $\approx$-closed relations $\overline{R}$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$.*

After this elimination of the exclusion atoms we still need to cope with $\subseteq_\approx$-atoms. Towards this end, let $\varphi \in \mathrm{FO}(\subseteq_\approx)$ and $\vartheta_1, \dots, \vartheta_k$ be an enumeration of all occurrences of inclusion atoms in φ, with $\vartheta_i := \overline{x_i} \subseteq_\approx \overline{y_i}$ for every $i \in \{1, \dots, k\}$. We shall use new relation symbols $R_{\vartheta_1}, \dots, R_{\vartheta_k}$ with the intended semantics that $R_{\vartheta_i} \subseteq X(\overline{y_i})_\approx$ where X is the team that "arrives" at ϑ_i. This will allow us to replace the subformulae ϑ_i by the formula $R_{\vartheta_i}\overline{x_i}$. However, this formula alone does not verify that $R_{\vartheta_i} \subseteq X(\overline{y_i})_\approx$ really holds. Additional formulae $\varphi^{(1)}(\overline{z_1}), \dots, \varphi^{(k)}(\overline{z_k})$ are required for the verification that values from R_{ϑ_i} could occur (up to equivalence) as a value for $\overline{y_i}$ in the team X that arrives at the corresponding inclusion atom. More precisely, $\varphi^{(i)}$ is constructed such that

$$(\mathfrak{A}, R_{\vartheta_1}, \dots, R_{\vartheta_k}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \varphi^{(i)}(\overline{z_i})$$

implies that the assignment s also satisfies φ and, more importantly, leads to an assignment s' that satisfies $s'(\overline{z_i}) \approx \overline{a}$ and that could be part of the team that satisfies the inclusion atom. Formally, we are going to prove that

$$\begin{aligned} \mathfrak{A} \models_X \varphi \Longleftrightarrow{} & \text{there are } \approx\text{-closed relations } R_{\vartheta_1}, \dots, R_{\vartheta_k} \text{ such that } (\mathfrak{A}, \overline{R}) \models_X \varphi^\star \text{ and} \\ & \text{for every } \overline{a} \in R_{\vartheta_i} \text{ there is an } s \in X \text{ with } (\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \varphi^{(i)}(\overline{z_i}). \end{aligned}$$

As already pointed out, $\varphi^\star$ results from φ by replacing every inclusion atom $\vartheta_i = \overline{x_i} \subseteq_\approx \overline{y_i}$ by $R_{\vartheta_i}\overline{x_i}$, while $\varphi^{(i)}$ is defined by induction (for every $i \in \{1, \dots, k\}$). Let ϑ be a subformula of φ. First-order literals are unchanged, i.e. $\vartheta^\star := \vartheta =: \vartheta^{(i)}$ if ϑ is such a literal. The inclusion atoms are translated as follows:

$$(\overline{x_j} \subseteq \overline{y_j})^{(i)} := \begin{cases} R_{\vartheta_i}\overline{x_i} \wedge \overline{y_i} \approx \overline{z_i}, & \text{if } i = j \\ R_{\vartheta_j}\overline{x_j}, & \text{if } i \neq j \end{cases}$$

Conjunctions and existential quantifiers are handled by defining

$$\begin{aligned} (\exists x\, \widetilde{\vartheta})^{(i)} &:= \exists x\, \widetilde{\vartheta}^{(i)} \text{ and} \\ (\widetilde{\vartheta}_1 \wedge \widetilde{\vartheta}_2)^{(i)} &:= \widetilde{\vartheta}_1^{(i)} \wedge \widetilde{\vartheta}_2^{(i)}. \end{aligned}$$

However, the translation of universal quantifiers or disjunctions is more complex:

$$\begin{aligned} (\widetilde{\vartheta}_1 \vee \widetilde{\vartheta}_2)^{(i)} &:= \begin{cases} \widetilde{\vartheta}_j^{(i)}, & \text{if } \overline{x_i} \subseteq_\approx \overline{y_i} \text{ occurs in } \widetilde{\vartheta}_j \\ (\widetilde{\vartheta}_1 \vee \widetilde{\vartheta}_2)^\star, & \text{otherwise} \end{cases} \\ (\forall x\, \widetilde{\vartheta})^{(i)} &:= \exists x\, \widetilde{\vartheta}^{(i)} \wedge (\forall x\, \widetilde{\vartheta})^\star. \end{aligned}$$

By construction we have that $(\forall x\, \vartheta)^\star$ is implied by $(\forall x\, \vartheta)^{(i)}$, because it is a subformula, while $\exists x\, \vartheta^{(i)}$ fetches the correct extension of the current assignment such that we end up with an assignment satisfying $\overline{y_i} \approx \overline{z_i}$ when arriving at the translation of $\overline{x_i} \subseteq_\approx \overline{y_i}$. The next lemma states that this construction actually captures the intuition that we have described above. The proof is given in the appendix.

▶ **Lemma 9.** *Let $\mathfrak{A} \in (\tau, \approx)$ and X be a team over $\mathfrak{A}$ with free(φ) = dom(X). Then the following are equivalent:*

(i) $\mathfrak{A} \models_X \varphi$

(ii) *There are $\approx$-closed relations $\overline{R} = (R_{\vartheta_1}, \dots, R_{\vartheta_k})$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists some $s \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \varphi^{(i)}$.*

We are now ready to show how inclusion atoms are translated into $\Sigma_1^1(\approx)$.

▶ **Theorem 10.** *For every formula $\varphi(\overline{x}) \in \mathrm{FO}(\subseteq_\approx)$ there exists a sentence $\varphi'(X) \in \Sigma^1_1(\approx)$ such that $\mathfrak{A} \models_X \varphi(\overline{x}) \iff (\mathfrak{A}, X) \models \varphi'(X)$ for every structure $\mathfrak{A}$ and every team X.*

Proof. Let

$$\varphi' := \exists_\approx R_{\vartheta_1} \dots \exists_\approx R_{\vartheta_k} \left(\forall \overline{x}(X\overline{x} \to \varphi^\star(\overline{x})) \wedge \bigwedge_{i=1}^{k} \forall \overline{z_i}(R_{\vartheta_i}\overline{z_i} \to \exists \overline{x}(X\overline{x} \wedge \varphi^{(i)}(\overline{x}, \overline{z_i}))) \right).$$

By construction, $(\mathfrak{A}, X) \models \varphi'$ if, and only if, there exist $\approx$-closed relations $\overline{R}$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_s \varphi^\star$ for every $s \in X$, and for every $\overline{a} \in R_i$ there exists some $s \in X$ with $(\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \varphi^{(i)}$. Since $\varphi^\star$ is a first-order formula, $\mathfrak{A} \models_s \varphi^\star$ for every $s \in X$ if, and only if, $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$. Hence, by Lemma 9, we can conclude that $(\mathfrak{A}, X) \models \varphi' \iff \mathfrak{A} \models_X \varphi$. ◀

In particular, every sentence $\varphi \in \mathrm{FO}(\subseteq_\approx)$ can be translated into an equivalent sentence $\varphi' \in \Sigma^1_1(\approx)$.

4 FO($\subseteq_\approx$) vs. GFP

An important result on logics with team semantics is the tight connection between inclusion logic and GFP$^+$, established by Galliani and Hella [6]. In this section we prove a similar result for FO($\subseteq_\approx$) by defining a fragment of GFP$^+$ which has the same expressive power as FO($\subseteq_\approx$).

▶ **Definition 11** (para-GFP$^+_\approx$)**.** The logic para-GFP$^+_\approx$ is defined as an extension of FO in negation normal form by the following formula formation rule. Let $k \geq 1$ and $\overline{R} = (R_1, \dots, R_k)$ be a tuple of unused relation symbols of arity $n_1, \dots, n_k$ respectively and let $(\varphi_i(\overline{R}, \overline{x_i}))_{i=1,\dots,k}$ be a tuple of $\mathrm{FO}(\tau \cup \{R_1, \dots, R_k\})$-formulae in negation normal form where $|\overline{x_i}| = n_i$ and every R_i occurs only positively in $\varphi_1, \dots, \varphi_k$. Furthermore, let $j \in \{1, \dots, k\}$ and $\overline{v}$ be a n_j-tuple of variables. Then

$$\varphi(\overline{v}) := [\text{para-GFP}_\approx\ (R_i, \overline{x_i})_{i=1,\dots,k} \,.\, (\varphi_i(\overline{R}, \overline{x_i}))_{i=1,\dots,k}]_j(\overline{v})$$

is a para-GFP$^+_\approx$-formula.

On every structure $\mathfrak{A} \in (\tau, \approx)$, the system $(\varphi_i(\overline{R}, \overline{x_i}))_{i=1,\dots,k}$ defines a parallel update operator $\Gamma^\mathfrak{A} : \mathcal{P}(A^{n_1}) \times \dots \times \mathcal{P}(A^{n_k}) \to \mathcal{P}(A^{n_1}) \times \dots \times \mathcal{P}(A^{n_k})$, by

$$\begin{aligned} \Gamma(\overline{R}) &:= \big(\Gamma_1(\overline{R}), \dots, \Gamma_k(\overline{R})\big) \quad \text{where} \\ \Gamma_i(\overline{R}) &:= [\![\varphi_i(\overline{R})]\!]^\mathfrak{A}_\approx = \{\overline{a} \in A^{n_i} \;:\; (\mathfrak{A}, \overline{R}) \models \varphi_i(\overline{R}, \overline{a})\}_\approx \end{aligned}$$

A tuple $(\mathfrak{A}, s)$ where $\mathfrak{A} \in (\tau, \approx)$ and $s : \{\overline{v}\} \to A$ is called a model of φ (and we write $\mathfrak{A} \models_s \varphi$ in this case) if, and only if, for the greatest fixed-point $\overline{S} = (S_1, \dots, S_k)$ of $\Gamma^\mathfrak{A}$ we have that $s(\overline{v}) \in S_j$.

The non-parallel variant GFP$^+_\approx$, where it is only allowed to use the operator para-GFP$_\approx$ in a non-parallel way, i.e. only in the following shape

$$[\mathrm{GFP}_\approx\ R\overline{x} \,.\, \varphi(R, \overline{x})](\overline{y}) := [\text{para-GFP}^+_\approx R\overline{x} \,.\, \varphi(R, \overline{x})]_1(\overline{y}),$$

has exactly the same expressive power as para-GFP$^+_\approx$.

The following lemma gives a characterization of the fixed-points of Γ:

▶ **Lemma 12** (Knaster-Tarski-Theorem for para-GFP$^+_\approx$). *Let*

$$\varphi(\overline{v}) = [\text{para-GFP}_\approx\ (R_i, \overline{x_i})_{i=1,\dots,k} \,.\, (\varphi_i(\overline{R}, \overline{x_i}))_{i=1,\dots,k}]_j(\overline{v})$$

be a para-GFP$^+_\approx$-formula, $\mathfrak{A} \in (\tau, \approx)$ and $\Gamma (= \Gamma^{\mathfrak{A}})$ be the corresponding parallel update operator w.r.t. $\varphi_1, \dots, \varphi_k$. For two given k-tuples $\overline{R}, \overline{S}$ of relations, we write $\overline{R} \subseteq \overline{S}$ if, and only if $R_i \subseteq S_i$ for every $i \in \{1, \dots, k\}$.

Let $X := \{\overline{S} : \overline{S} \subseteq \Gamma(\overline{S})\}$. Then $\bigcup X := (Y_1, \dots, Y_k)$ where for every $j \in \{1, \dots, k\}$, $Y_j := \bigcup_{\overline{S} \in X} S_j$ is the greatest fixed-point of Γ. Furthermore, these Y_j are $\approx$-closed.

4.1 From FO($\subseteq_\approx$) to GFP$^+_\approx$

▶ **Theorem 13.** *For every formula $\varphi(\overline{x}) \in \text{FO}(\subseteq_\approx)$ there exists a sentence $\varphi^+ \in \text{GFP}^+_\approx$ such that $\mathfrak{A} \models_X \varphi \iff (\mathfrak{A}, X) \models \varphi^+$ for every structure $\mathfrak{A} \in (\tau, \approx)$ and every team X over $\mathfrak{A}$.*

Proof. In the last section we have presented the FO-formulae $\varphi^\star(\overline{R})$ and $\varphi^{(i)}(\overline{R})$ (for $i \in \{1, \dots, k\}$) using new relation symbols $\overline{R} = (R_1, \dots, R_k)$ such that for every $\mathfrak{A} \in (\tau, \approx)$ and every team X over $\mathfrak{A}$ with $\text{dom}(X) \supseteq \text{free}(\varphi)$ the following are equivalent:

(1) $\mathfrak{A} \models_X \varphi$

(2) There are $\approx$-closed relations $\overline{R}$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_i$ there exists some $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \varphi^{(i)}$.

Furthermore, the relation symbols $R_1, \dots, R_k$ occur only positively in $\varphi^\star$ and $\varphi^{(i)}$ and the tuple $\overline{z_i}$ occurs exactly once in a subformula of the form $\overline{x_i} \approx \overline{z_i}$ in $\varphi^{(i)}$. Let $\widetilde{\varphi}^\star$ and the $\widetilde{\varphi}^{(i)}$ be the formulae that result from $\varphi^\star, \varphi^{(i)}$ by replacing every occurrence of the form $R_i\overline{v}$ by its guarded version $(R_i)_\approx \overline{v} := \exists \overline{w}(\overline{v} \approx \overline{w} \land R_i\overline{w})$. This allows us to drop the requirement that the relations $\overline{R}$ are $\approx$-closed.

▶ Claim 14. For every $\mathfrak{A}$ and every team X over $\mathfrak{A}$, (1) and (2) are equivalent to:

(3) There are relations $\overline{R}$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_X \widetilde{\varphi}^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_i$ there exists some $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \widetilde{\varphi}^{(i)}$.

To prove this, one has to exploit the fact that every R_j ($j \in \{1, \dots, k\}$) occurs only $\approx$-guarded in $\widetilde{\varphi}^\star, \widetilde{\varphi}^{(1)}, \dots, \widetilde{\varphi}^{(k)}$ and the variables $\overline{z_i}$ occur (exactly once) in a subformula of the form $\overline{w} \approx \overline{z_i}$ in $\varphi^{(i)}$. By expressing (3) in existential second-order logic, we obtain the following equivalent statement:

(4) $(\mathfrak{A}, X) \models \exists \overline{R}\left(\forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^\star(\overline{R}, \overline{x})) \land \psi\right)$ where $\psi := \bigwedge_{i=1}^k \forall \overline{z_i}(R_i\overline{z_i} \to \eta_i(\overline{R}, \overline{z_i}))$ and $\eta_i(\overline{R}, \overline{z_i}) := \exists \overline{x}(X\overline{x} \land \widetilde{\varphi}^{(i)}(\overline{R}, \overline{z_i}, \overline{x}))$.

Let $\Gamma(\overline{R}) := \left(\Gamma_1(\overline{R}), \dots, \Gamma_k(\overline{R})\right)$ where

$$\Gamma_i(\overline{R}) := [\![\eta_i(\overline{R}, \overline{z_i})]\!]^{(\mathfrak{A}, X)} = \{\overline{a} \in A^{\text{ar}(R_i)} : (\mathfrak{A}, X(\overline{x}), \overline{R}) \models \eta_i(\overline{a})\}.$$

Note that $[\![\eta_i(\overline{R}, \overline{z_i})]\!]^{(\mathfrak{A}, X(\overline{x}))} = [\![\eta_i(\overline{R}, \overline{z_i})]\!]_\approx^{(\mathfrak{A}, X(\overline{x}))}$, because the free variables $\overline{z_i}$ occur exactly once in a subformula of the form $\overline{w} \approx \overline{z_i}$. This is the reason why Γ is the para-GFP$^+_\approx$-update operator w.r.t. $\eta_1, \dots, \eta_k$.

Furthermore, $(\mathfrak{A}, X, \overline{R}) \models \forall \overline{z_i}(R_i\overline{z_i} \to \eta_i(\overline{R}, \overline{z_i}))$ if, and only if, $R_i \subseteq \Gamma_i(\overline{R})$. Consequently, we have $(\mathfrak{A}, X, \overline{R}) \models \psi$ if, and only if, $\overline{R} \subseteq \Gamma(\overline{R})$.

▶ Claim 15. For $j \leq k$, let $\vartheta_j(\overline{z_j}) := [\text{para-GFP}_\approx\ (R_i, \overline{z_i})_{i=1,\dots,k} \,.\, (\eta_i(\overline{R}, \overline{z_i}))_{i=1,\dots,k})]_j(\overline{z_j})$, and let γ result from $\widetilde{\varphi}^\star$ by replacing every occurrence of $R_j(\overline{w})$ by $\vartheta_j(\overline{w})$. Then, for every $\mathfrak{A} \in (\tau, \approx)$ and every team X, (4) is equivalent to

(5) $(\mathfrak{A}, X) \models \forall \overline{x}(X\overline{x} \to \gamma)$.

Proof. (4) $\Longrightarrow$ (5) : Let $(\mathfrak{A}, X) \models \exists \overline{R}\left(\forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^{\star}(\overline{R}, \overline{x})) \wedge \psi\right)$. Then there are relations $\overline{R}$ such that $(\mathfrak{A}, X, \overline{R}) \models \forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^{\star}(\overline{R}, \overline{x}))$ and $(\mathfrak{A}, X, \overline{R}) \models \psi$. As observed above, it follows that $\overline{R} \subseteq \Gamma(\overline{R})$. So, by Lemma 12, $\overline{R} \subseteq \overline{S}$ where $\overline{S}$ is the greatest fixed-point of Γ. Since we have $(\mathfrak{A}, X, \overline{R}) \models \forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^{\star}(\overline{R}, \overline{x}))$ and the relations symbols $R_1, \dots, R_k$ occur only positively in $\widetilde{\varphi}^{\star}$, we can conclude that $(\mathfrak{A}, X, \overline{S}) \models \forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^{\star}(\overline{S}, \overline{x}))$. Because $\overline{S}$ is the greatest fixed-point of Γ, it follows that $S_i = [\![\vartheta_i(\overline{z_i})]\!]^{(\mathfrak{A},X)}$ and, by construction of γ, we obtain that $(\mathfrak{A}, X) \models \forall \overline{x}(X\overline{x} \to \gamma)$.

(5) $\Longrightarrow$ (4) : Let $(\mathfrak{A}, X) \models \forall \overline{x}(X\overline{x} \to \gamma)$ and let $\overline{S}$ be the greatest fixed-point of Γ. Then $(\mathfrak{A}, X, \overline{S}) \models \forall \overline{x}(X\overline{x} \to \widetilde{\varphi}^{\star})$ and $\overline{S} = \Gamma(\overline{S})$. Therefore, we have $(\mathfrak{A}, X, \overline{S}) \models \forall \overline{z_i}(S_i\overline{z_i} \to \eta_i(\overline{z_i}))$ for every $i \in \{1, \dots, k\}$ and, hence, $(\mathfrak{A}, X, \overline{S}) \models \psi(\overline{S})$. ◀

By construction $\forall \overline{x}(X\overline{x} \to \gamma) \in$ para-GFP$^+_{\approx}$. Since para-GFP$^+_{\approx}$ has the same expressive power as GFP$^+_{\approx}$, there also exists a sentence $\varphi^+ \in$ GFP$^+_{\approx}$ that is equivalent to $\varphi(\overline{x})$. ◀

4.2 From GFP$^+_{\approx}$ to FO($\subseteq_{\approx}$)

In order to translate a given sentence $\varphi \in$ GFP$^+_{\approx}$ into a FO($\subseteq_{\approx}$)-formula, we assume that φ is in a normal form which is given by the following lemma. By using adaptations of ideas from [6] we then show that such a sentence can be expressed in FO($\subseteq_{\approx}$).

▶ **Lemma 16.** *For every sentence $\varphi \in$ para-GFP$^+_{\approx}$ there exists a formula $\psi(R, \overline{x}) \in$ FO, in which R occurs only positively and only $\approx$-guarded, such that φ is equivalent to*

$$\exists \overline{v}[GFP_{\approx}\ R\overline{x}\,.\,\psi(R, \overline{x})](\overline{v}).$$

Our next lemma shows that we can eliminate the relation symbol R in ψ by introducing $\subseteq_{\approx}$-atoms and encoding R in a tuple $\overline{x}$ of variables.

▶ **Lemma 17.** *Let R be a relation symbol of arity n, let $\overline{x}, \overline{y}$ be tuples of variables where $|\overline{x}| = n$ (whereas $\overline{y}$ is of arbitrary length and can also be empty). Furthermore, let $\psi(R, \overline{x}, \overline{y}) \in$ FO$(\tau \cup \{R\})$ be a first-order formula in which R occurs only positively and $\approx$-guarded, and with free$(\psi) \subseteq \{\overline{x}, \overline{y}\}$ such that the variables in $\overline{x}$ are never quantified in ψ. Then there exists a formula $\psi^{\star}(\overline{x}, \overline{y}) \in$ FO($\subseteq_{\approx}$) of signature τ such that for every $\mathfrak{A} \in (\tau, \approx)$ and every team X we have that*

$$\mathfrak{A} \models_X \psi^{\star}(\overline{x}, \overline{y}) \Longleftrightarrow (\mathfrak{A}, X(\overline{x})) \models_s \psi(R, \overline{x}, \overline{y}) \text{ for every } s \in X.$$

This lemma can be shown by induction over the structure of ψ. Now we are able to express [GFP$_{\approx}$ $R\overline{x}\,.\,\varphi(R, \overline{x})$] in FO($\subseteq_{\approx}$).

▶ **Theorem 18.** *Let $\psi(R, \overline{x}) \in$ FO where $ar(R) = |\overline{x}|$, R occurs only positively and $\approx$-guarded in ψ, and the variables in $\overline{x}$ are never quantified in ψ. Then there exists a formula $\psi^+(\overline{x}) \in$ FO($\subseteq_{\approx}$) such that for every $\mathfrak{A} \in (\tau, \approx)$ and every team X we have that*

$$\mathfrak{A} \models_X \psi^+(\overline{x}) \Longleftrightarrow \mathfrak{A} \models_s [GFP_{\approx}\ R\overline{x}\,.\,\psi(R, \overline{x})](\overline{x}) \text{ for every } s \in X.$$

Proof. Let $\psi^+(\overline{x}) \coloneqq \exists \overline{y}(\overline{x} \subseteq_{\approx} \overline{y} \wedge \exists \overline{z}(\overline{y} \approx \overline{z} \wedge \psi^{\star}(\overline{z})))$.

"$\Longrightarrow$": First we assume that $\mathfrak{A} \models_X \psi^+(\overline{x})$. Then there exists a function $F : X \to \mathcal{P}(A^n) \setminus \{\varnothing\}$ such that $\mathfrak{A} \models_Y \overline{x} \subseteq_{\approx} \overline{y} \wedge \exists \overline{z}(\overline{y} \approx \overline{z} \wedge \psi^{\star}(\overline{z}))$ where $Y \coloneqq X[\overline{y} \mapsto F]$. So there exists a function $G : Y \to \mathcal{P}(A^n) \setminus \{\varnothing\}$ satisfying $\mathfrak{A} \models_Z \overline{y} \approx \overline{z} \wedge \psi^{\star}(\overline{z})$ where $Z \coloneqq Y[\overline{z} \mapsto G]$. By Lemma 17, it follows that

$$(\mathfrak{A}, Z(\overline{z})) \models_s \psi(R, \overline{z}) \text{ for every } s \in Z.$$

So we have $Z(\overline{z}) \subseteq [\![\psi(R,\overline{z})]\!]^{(\mathfrak{A},Z(\overline{z}))} \subseteq [\![\psi(R,\overline{z})]\!]_{\approx}^{(\mathfrak{A},Z(\overline{z}))} = \Gamma_\psi(Z(\overline{z}))$ where $\Gamma_\psi \coloneqq \Gamma_\psi^{\mathfrak{A}}$ is the $\mathrm{GFP}_{\approx}^{+}$-update operator w.r.t. ψ. It follows that $Z(\overline{z}) \subseteq \mathrm{gfp}(\Gamma_\psi)$ (by Lemma 12). Since $\mathrm{gfp}(\Gamma_\psi)$ is $\approx$-closed and $X(\overline{x}) \subseteq_{\approx} Y(\overline{y}) \approx Z(\overline{z})$, we have that $X(\overline{x}) \subseteq \mathrm{gfp}(\Gamma_\psi)$. Hence, we obtain that $\mathfrak{A} \models_s [\mathrm{GFP}_{\approx}\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{x})$ for every $s \in X$.

"$\Longleftarrow$": Now we assume that $\mathfrak{A} \models_s [\mathrm{GFP}_{\approx}\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{x})$ for every $s \in X$. If $X = \varnothing$, then $\mathfrak{A} \models_X \psi^+(\overline{x})$ follows from the empty team property. Henceforth, let $X \neq \varnothing$. Let $\Gamma_\psi = \Gamma_\psi^{\mathfrak{A}}$ be the $\mathrm{GFP}_{\approx}^{+}$-update operator defined w.r.t. $\psi(R)$. From our assumption follows that $X(\overline{x}) \subseteq \mathrm{gfp}(\Gamma_\psi)$. Since $X \neq \varnothing$, it follows that $\mathrm{gfp}(\Gamma_\psi) \neq \varnothing$. Our goal is to prove that $\mathfrak{A} \models_X \psi^+(\overline{x})$. Towards this end, we define $F : X \to \mathcal{P}(A^n) \setminus \{\varnothing\}$, $F(s) \coloneqq \mathrm{gfp}(\Gamma_\psi)$ and $Y \coloneqq X[\overline{y} \mapsto F]$ and claim that $\mathfrak{A} \models_Y \overline{x} \subseteq_{\approx} \overline{y} \wedge \exists \overline{z}(\overline{y} \approx \overline{z} \wedge \psi^\star(\overline{z}))$. Since $Y(\overline{x}) = X(\overline{x}) \subseteq \mathrm{gfp}(\Gamma_\psi) = Y(\overline{y})$ it is clear that $\mathfrak{A} \models_Y \overline{x} \subseteq_{\approx} \overline{y}$.

We still need to prove that $\mathfrak{A} \models_Y \exists \overline{z}(\overline{y} \approx \overline{z} \wedge \psi^\star(\overline{z}))$. By definition of Y, we know that $Y(\overline{y}) = \mathrm{gfp}(\Gamma_\psi) = \Gamma_\psi(\mathrm{gfp}(\Gamma_\psi)) = [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]_{\approx}^{\mathfrak{A}}$. This implies that for every $s \in Y$ there exists some $\overline{a} \in [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}}$ such that $\overline{a} \approx s(\overline{y})$.

Let $G : Y \to \mathcal{P}(A^n) \setminus \{\varnothing\}$ be given by

$$G(s) \coloneqq \{\overline{a} \in [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}} \;:\; s(\overline{y}) \approx \overline{a}\}$$

and $Z \coloneqq Y[\overline{z} \mapsto G]$. Clearly it holds that $Z(\overline{z}) \subseteq [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}}$. We claim that even $Z(\overline{z}) = [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}}$ is true. To see this, let $\overline{a} \in [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}} \subseteq [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]_{\approx}^{\mathfrak{A}} = Y(\overline{y})$. So there exists an $s \in Y$ with $s(\overline{y}) \approx \overline{a}$. Hence, we have that $\overline{a} \in G(s)$ and, consequently, $\overline{a} \in Z(\overline{z})$.

It is the case that $\mathfrak{A} \models_Z \overline{y} \approx \overline{z}$, because this follows from the definition of G. Now we prove that $\mathfrak{A} \models_Z \psi^\star(\overline{z})$. By Lemma 17, we need to verify that $(\mathfrak{A}, Z(\overline{z})) \models_s \psi(R,\overline{z})$ for every $s \in Z$. In other words, we need to verify that $Z(\overline{z}) \subseteq [\![\psi(Z(\overline{z}),\overline{z})]\!]^{\mathfrak{A}}$. Since $Z(\overline{z}) = [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}}$, we can conclude that

$$[\![\psi(Z(\overline{z}),\overline{z})]\!]^{\mathfrak{A}} = [\![\psi([\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}},\overline{z})]\!]^{\mathfrak{A}}$$

Due to the fact that R occurs only $\approx$-guarded in ψ, we can observe that

$$\begin{aligned}[\![\psi([\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]^{\mathfrak{A}},\overline{z})]\!]^{\mathfrak{A}} &= [\![\psi([\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{x})]\!]_{\approx}^{\mathfrak{A}},\overline{z})]\!]^{\mathfrak{A}} \\ &= [\![\psi(\Gamma_\psi(\mathrm{gfp}(\Gamma_\psi)),\overline{z})]\!]^{\mathfrak{A}} \\ &= [\![\psi(\mathrm{gfp}(\Gamma_\psi),\overline{z})]\!]^{\mathfrak{A}} = Z(\overline{z})\end{aligned}$$

Therefore, we have $Z(\overline{z}) = [\![\psi(Z(\overline{z}),\overline{z})]\!]^{\mathfrak{A}}$ which implies that $Z(\overline{z}) \subseteq [\![\psi(Z(\overline{z}),\overline{z})]\!]^{\mathfrak{A}}$. So we have $(\mathfrak{A}, Z(\overline{z})) \models_s \psi(R,\overline{z})$ for every $s \in Z$, which concludes the proof of $\mathfrak{A} \models_Z \psi^\star$ and of $\mathfrak{A} \models_X \psi^+$. ◀

▶ **Corollary 19.** *For every $\mathrm{GFP}_{\approx}^{+}$-sentence φ there is an equivalent sentence $\vartheta \in \mathrm{FO}(\subseteq_{\approx})$.*

Proof. Let $\varphi \in \mathrm{GFP}_{\approx}^{+}$. By Lemma 16, there exists a first-order formula $\psi(R,\overline{x})$ where the n-ary relation symbol R occurs only positively and only $\approx$-guarded in ψ such that

$$\varphi \equiv \exists \overline{v}[\mathrm{GFP}_{\approx}\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{v}).$$

W.l.o.g. we can assume that the variables in $\overline{x}$ are never quantified in ψ. So, by Theorem 18, it follows that there exists some $\psi^+(\overline{x}) \in \mathrm{FO}(\subseteq_{\approx})$ such that for every $\mathfrak{A} \in (\tau, \approx)$ and every team X over $\mathfrak{A}$ with $\mathrm{dom}(X) \supseteq \{\overline{x}\}$ holds

$$\mathfrak{A} \models_X \psi^+(\overline{x}) \Longleftrightarrow \mathfrak{A} \models_s [\mathrm{GFP}_{\approx}\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{x}) \text{ for every } s \in X$$

Let $\vartheta := \exists\overline{v}\psi^+(\overline{v})$ and $\mathfrak{A} \in (\tau, \approx)$. Our goal is to prove that $\mathfrak{A} \models \varphi \Longleftrightarrow \mathfrak{A} \models \vartheta$.

"$\Longleftarrow$": Let $\mathfrak{A} \models \vartheta$. Then there exists a function $F : \{\varnothing\} \to \mathcal{P}(A^{|\overline{v}|}) \setminus \{\varnothing\}$ such that $\mathfrak{A} \models_Y \psi^+(\overline{v})$ where $Y = \{\varnothing\}[\overline{v} \mapsto F]$. Then we have $\mathfrak{A} \models_s [\mathrm{GFP}_\approx\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{v})$ for every $s \in Y$ and, since Y is non-empty, it follows that $\mathfrak{A} \models \exists\overline{v}[\mathrm{GFP}_\approx\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{v})$.

"$\Longrightarrow$": Now let $\mathfrak{A} \models \varphi \equiv \exists\overline{v}[\mathrm{GFP}_\approx\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{v})$. Then there exists some $\overline{a} \in A$ such that $\mathfrak{A} \models [\mathrm{GFP}_\approx\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{a})$. Let $Y = \{s\}$ be the singleton team consisting only of s with $s(\overline{v}) = \overline{a}$. Then it follows that $\mathfrak{A} \models_s [\mathrm{GFP}_\approx\, R\overline{x}\,.\,\psi(R,\overline{x})](\overline{v})$ for every $s \in Y$ and, consequently, $\mathfrak{A} \models_Y \psi^+(\overline{v})$, proving that $\mathfrak{A} \models_{\{\varnothing\}} \exists\overline{v}\psi^+(\overline{v}) = \vartheta$. ◀

5 $\Sigma^1_1(\approx)$ on restricted classes of structures

In this section we compare $\Sigma^1_1(\approx)$ with FO and Σ^1_1 and study how restrictions imposed on the given equivalence influence the expressive power of $\Sigma^1_1(\approx)$. Our first result is that the expressive power of $\Sigma^1_1(\approx) \equiv \mathrm{FO}(\subseteq_\approx, |_\approx)$ lies strictly between FO and Σ^1_1. Furthermore, we also have $\mathrm{FO} < \mathrm{FO}(\subseteq_\approx, |_\approx) < \Sigma^1_1$ on the class of structures with only a bounded number of non-trivial equivalence classes and on the class of structures where each equivalence class is of size $\leq k$ (for some fixed $k > 1$). However, when restricting both the size of the equivalence classes and the number of non-trivial equivalence classes, then $\mathrm{FO}(\subseteq_\approx, |_\approx)$ has the same expressive power as Σ^1_1. To prove these results, we use an adaption of the Ehrenfeucht-Fraïssé method for $\mathrm{FO}(\subseteq_\approx, |_\approx)$, which relies on the games presented in [15].

▶ **Definition 20.** Let $\mathfrak{A}, \mathfrak{B} \in (\tau, \approx)$, $n \in \mathbb{N}$ and $\Omega_\approx \subseteq \{\mathrm{dep}_\approx, \perp^\approx, \subseteq_\approx, |_\approx\}$. The game $\mathcal{G}_{\Omega_\approx,n}(\mathfrak{A}, \mathfrak{B})$ is played by two players which are called Duplicator and Spoiler. The positions of the game are tuples (X, Y) of teams over $\mathfrak{A}, \mathfrak{B}$ with $\mathrm{dom}(X) = \mathrm{dom}(Y)$. Unless stated otherwise the game starts at position $(\{\varnothing\}, \{\varnothing\})$ and then n moves are played. In each move Spoiler always chooses between one of the following 3 moves to continue the game:

1. Move $\vee$:
   - Spoiler represents X as a union $X = X_0 \cup X_1$.
   - Duplicator replies with a representation of Y as $Y = Y_0 \cup Y_1$.
   - Spoiler chooses $i \in \{0, 1\}$ and the game continues at position (X_i, Y_i).
2. Move $\exists$:
   - Spoiler chooses a function $F : X \to \mathcal{P}(A) \setminus \{\varnothing\}$.
   - Duplicator replies with a function $G : Y \to \mathcal{P}(B) \setminus \{\varnothing\}$.
   - The game continues at position $(X[v \mapsto F], Y[v \mapsto G])$ where v is a new variable.
3. Move $\forall$:
   - The game continues at position $(X[v \mapsto A], Y[v \mapsto B])$ where v is a new variable.

Positions (X, Y) with $\mathfrak{A} \models_X \vartheta$ but $\mathfrak{B} \not\models_Y \vartheta$ for some literal $\vartheta \in \mathrm{FO}(\Omega_\approx)$ are Spoiler's winning position. Duplicator wins, if such positions are avoided for n moves.

The game $\mathcal{G}_{\Omega_\approx}(\mathfrak{A}, \mathfrak{B})$ is played similarly: first Spoiler chooses a number $n \in \mathbb{N}$ and then $\mathcal{G}_{\Omega_\approx,n}(\mathfrak{A}, \mathfrak{B})$ is played.

These games characterize semi-equivalences of $\mathfrak{A}$ and $\mathfrak{B}$ (up to a certain depth). The depth of $\varphi \in \mathrm{FO}(\Omega_\approx)$, denoted as $\mathrm{depth}(\varphi)$, is defined inductively:

$$
\begin{aligned}
\mathrm{depth}(\vartheta) &:= 0 \text{ for every literal } \vartheta \in \mathrm{FO}(\Omega_\approx) \\
\mathrm{depth}(\exists v\varphi') &:= \mathrm{depth}(\varphi') + 1 =: \mathrm{depth}(\forall v\varphi') \\
\mathrm{depth}(\varphi_1 \vee \varphi_2) &:= \max(\mathrm{depth}(\varphi_1), \mathrm{depth}(\varphi_2)) + 1 \\
\mathrm{depth}(\varphi_1 \wedge \varphi_2) &:= \max(\mathrm{depth}(\varphi_1), \mathrm{depth}(\varphi_2))
\end{aligned}
$$

▶ **Definition 21** (Semi-equivalence, [15]). Let $\mathfrak{A}, \mathfrak{B} \in (\tau, \approx)$ and X, Y be teams over $\mathfrak{A}, \mathfrak{B}$ with $\mathrm{dom}(X) = \mathrm{dom}(Y)$. We write $\mathfrak{A}, X \Rrightarrow_{\Omega_\approx, n} \mathfrak{B}, Y$ (and say that $\mathfrak{A}, X$ is semi-equivalent to $\mathfrak{B}, Y$ up to depth n), if $\mathfrak{A} \models_X \varphi$ implies $\mathfrak{B} \models_Y \varphi$ for every $\varphi \in \mathrm{FO}(\Omega_\approx)$ with $\mathrm{depth}(\varphi) \leq n$. Furthermore, we write $\mathfrak{A}, X \Rrightarrow_{\Omega_\approx} \mathfrak{B}, Y$, if $\mathfrak{A}, X \Rrightarrow_{\Omega_\approx, n} \mathfrak{B}, Y$ for every $n \in \mathbb{N}$. When $\Omega_\approx$ is clear from the context, we sometimes omit it as a subscript.

In first-order logic, the concept of semi-equivalence coincides with the usual equivalence concept between structures, but this is not the case in logics with team semantics. For example $\mathfrak{A}, X \Rrightarrow \mathfrak{B}, \varnothing$ follows from the empty team property, but $\mathfrak{B}, \varnothing \Rrightarrow \mathfrak{A}, X$ is not true in general. We write $\mathfrak{A}, X \equiv_n \mathfrak{B}, Y$, if $\mathfrak{A}, X \Rrightarrow_n \mathfrak{B}, Y$ and $\mathfrak{B}, Y \Rrightarrow_n \mathfrak{A}, X$. $\mathfrak{A}, X \equiv \mathfrak{B}, Y$ is defined analogously.

▶ **Theorem 22.** *Let τ be a finite signature and $\mathfrak{A}, \mathfrak{B} \in (\tau, \approx)$. Duplicator has a winning strategy for $\mathcal{G}_{\Omega_\approx, n}(\mathfrak{A}, \mathfrak{B})$ from position (X, Y) if, and only if $\mathfrak{A}, X \Rrightarrow_{\Omega_\approx, n} \mathfrak{B}, Y$.*

Having these games at our disposal, we can prove that $\mathrm{FO}(\subseteq_\approx, |_\approx)$ is strictly less powerful than Σ^1_1. Consider the following problem:

$$\mathcal{C}_{\mathrm{even}} := \{\mathfrak{A} \in (\tau, \approx) \,:\, \text{there is some } a \in A \text{ such that } |[a]_\approx| \text{ is even}\}.$$

▶ **Theorem 23.** *$\mathcal{C}_{even}$ is not expressible in $\mathrm{FO}(\subseteq_\approx, |_\approx)$.*

We just give a short sketch of the proof: Consider $\mathfrak{A}_m := (A_m, \approx^{\mathfrak{A}_m})$ and $\mathfrak{B}_m := (B_m, \approx^{\mathfrak{B}_m})$ where $|A_m| = 2m$, $|B_m| = 2m + 1$, $\approx^{\mathfrak{A}_m} := A_m \times A_m$ and $\approx^{\mathfrak{B}_m} := B_m \times B_m$. Then $\mathfrak{A}_m \in \mathcal{C}_{\mathrm{even}}$ while $\mathfrak{B}_m \notin \mathcal{C}_{\mathrm{even}}$. It is not difficult to prove that Duplicator wins the games $\mathcal{G}_m(\mathfrak{A}_m, \mathfrak{B}_m)$ and $\mathcal{G}_m(\mathfrak{B}_m, \mathfrak{A}_m)$ by maintaining as an invariant that the equality types induced by the assignments in the two teams are always equal. On the other hand, it is easy to see that $\mathrm{FO}(\mathrm{dep}_\approx)(\leq \mathrm{FO}(\subseteq_\approx, |_\approx))$ can express that the number of equivalence classes is even, but this is not definable in first-order logic.

▶ **Corollary 24.** $\mathrm{FO} < \mathrm{FO}(\subseteq_\approx, |_\approx) < \Sigma^1_1$.

Next we study whether restrictions imposed on the given equivalence influence the expressive power of Σ^1_1. Consider the class $\mathcal{K}_{\leq p}$ of structures $\mathfrak{A} \in (\tau, \approx)$ where every equivalence class of $\mathfrak{A}$ is of size $\leq p$. On $\mathcal{K}_{\leq 1}$, $\Sigma^1_1(\approx)$ has the same expressive power as Σ^1_1, because every relation over $\mathfrak{A} \in \mathcal{K}_{\leq 1}$ is $\approx^{\mathfrak{A}}$-closed. However, this is not the case for $p \geq 2$ as the next theorem shows.

▶ **Theorem 25.** *Let $p \geq 2$. $\mathrm{FO} < \mathrm{FO}(\subseteq_\approx, |_\approx) < \Sigma^1_1$ holds on the class $\mathcal{K}_{\leq p}$ of structures $\mathfrak{A} \in (\tau, \approx)$ with $|[a]_\approx| \leq p$ for every $a \in A$.*

To prove this (see appendix), we are using an Ehrenfeucht-Fraïssé argument and prove that $\mathrm{FO}(\subseteq_\approx, |_\approx)$ is unable to express non-connectivity of graphs when the equivalence classes are allowed to contain up to 2 elements.

Restricting the number of equivalence classes is not really interesting, because it leads to a situation where $\Sigma^1_1(\approx)$ has the same expressive power as FO, because there are only $2^{(k^r)}$ many $\approx$-closed relations of arity r when k is the number of $\approx$-classes, which can be simulated in first-order logic.

Another possible restriction is to admit only a bounded number of non-trivial equivalence classes (which consist of more than one element). Let $\mathcal{K}_{\mathrm{NT}\leq p}$ be the class of all $\mathfrak{A} \in (\tau, \approx)$ with at most p many non-trivial equivalence classes (for some $p \geq 1$).

But then again, $\mathcal{C}_{\mathrm{even}} \cap \mathcal{K}_{\mathrm{NT}\leq p}$ is not definable in $\mathrm{FO}(\subseteq_\approx, |_\approx)$ on $\mathcal{K}_{\mathrm{NT}\leq p}$. Hence, we also have $\mathrm{FO} < \Sigma^1_1(\approx) < \Sigma^1_1$ on $\mathcal{K}_{\mathrm{NT}\leq p}$.

However, combining the conditions imposed on the number of non-trivial equivalence and their size, leads to an interesting situation: $\Sigma^1_1(\approx) \equiv \Sigma^1_1$ on $\mathcal{K}_{\mathrm{NT}\leq p_1, \leq p_2} := \mathcal{K}_{\mathrm{NT}\leq p_1} \cap \mathcal{K}_{\leq p_2}$. The reason for this is that at most $p_1 \cdot p_2$ many elements are located inside non-trivial equivalence classes, while all the other elements are only equivalent to themselves. Since $\Sigma^1_1(\approx)$ allows us to obtain a linear order on the equivalence classes, it is possible to encode arbitrary relations and, hence, to simulate Σ^1_1.

References

1. A. Blass and Y. Gurevich. Henkin quantifiers and complete problems. *Annals of Pure and Applied Logic*, 32:1–16, 1986.
2. H. Enderton. Finite partially ordered quantifiers. *Z. Math. Logik*, 16:393–397, 1970.
3. F. Engström. Generalized quantifiers in dependence logic. *Journal of Logic, Language, and Information*, 2012.
4. S. Abramsky et al., editor. *Dependence Logic. Theory and Applications*. Birkhäuser, 2016.
5. P. Galliani. Inclusion and exclusion in team semantics — on some logics of imperfect information. *Annals of Pure and Applied Logic*, 163:68–84, 2012.
6. P. Galliani and L. Hella. Inclusion Logic and Fixed Point Logic. In *Computer Science Logic 2013*, pages 281–295, 2013.
7. E. Grädel. Games for inclusion logic and fixed-point logic. In S. Abramsky et al., editor, *Dependence Logic. Theory and Applications*. Birkhäuser, 2016.
8. E. Grädel and J. Väänänen. Dependence and independence. *Studia Logica*, 101(2):399–410, 2013.
9. L. Henkin. Some remarks on infinitely long formulas. *Journal of Symbolic Logic*, pages 167–183, 1961.
10. J. Hintikka and G. Sandu. Informational independence as a semantical phenomenon. In *Studies in Logic and Foundations of Mathematics*, volume 126, pages 571–589. North-Holland, 1989.
11. W. Hodges. Compositional semantics for a logic of imperfect information. *Logic Journal of IGPL*, 5:539–563, 1997.
12. J. Kontinen and J. Väänänen. On definability in dependence logic. *Journal of Logic, Language, and Information*, 18:317–241, 2009.
13. A. Mann, G. Sandu, and M. Sevenster. *Independence-Friendly Logic. A Game-Theoretic Approach*, volume 386 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, 2012.
14. R. Rönnholm. Capturing k-ary existential second-order logic with k-ary inclusion-exclusion logic. *Ann. Pure Appl. Logic*, 169(3):177–215, 2018.
15. J. Väänänen. *Dependence logic: A new approach to independence friendly logic*, volume 70. Cambridge University Press, 2007.
16. W. Walkoe. Finite partially-ordered quantification. *Journal of Symbolic Logic*, 35:535–555, 1970.

A Appendix

A.1 The Expressive Power of $\mathrm{FO}(\subseteq_\approx, |_\approx)$

Proof of Theorem 8. By induction:

Case $\varphi = \overline{v} \mid_\approx \overline{w}$: Then $\varphi^\star := R_\varphi\overline{v} \wedge \neg R_\varphi\overline{w}$. Let $R_\varphi := (X(\overline{v}))_\approx$, which is the $\approx$-closure of $X(\overline{v})$. Now we observe that

$$\begin{aligned}\mathfrak{A} \models_X \overline{v} \mid_\approx \overline{w} &\iff \text{for every } s, s' \in X: \ s(\overline{v}) \not\approx s'(\overline{w}) \\ &\iff \text{for every } s \in X: \ s(\overline{v}) \in R_\varphi \text{ and } s(\overline{w}) \notin R_\varphi \\ &\iff (\mathfrak{A}, R_\varphi) \models_X R_\varphi\overline{v} \wedge \neg R_\varphi\overline{w}.\end{aligned}$$

Case φ is some FO-literal or an $\subseteq_\approx$-atom: Then we have $\varphi^\star = \varphi$ and there is nothing to prove, because the relation symbols $\overline{R}$ do not occur in $\varphi^\star$.

Case $\varphi = \varphi_0 \vee \varphi_1$: Let $\vartheta_1^{(j)}, \dots, \vartheta_{k_j}^{(j)}$ be the exclusion atoms that occur in φ_j.

"$(i) \Longrightarrow (ii)$": First, we assume that $\mathfrak{A} \models_X \varphi_0 \vee \varphi_1$. Then there are teams X_0, X_1 such that $X = X_0 \cup X_1$ and $\mathfrak{A} \models_{X_j} \varphi_j$ for $j \in \{0,1\}$. By induction hypothesis, there exists two tuples of $\approx$-closed relations $\overline{R_j} = (R_{\vartheta_1^{(j)}}, \dots, R_{\vartheta_{k_j}^{(j)}})$ such that $(\mathfrak{A}, \overline{R_j}) \models_{X_j} \varphi_j^\star$ (for $j \in \{0,1\}$). We define

$$\overline{R} := (R_{\vartheta_1^{(0)}}, \dots, R_{\vartheta_{k_0}^{(0)}}, R_{\vartheta_1^{(1)}}, \dots, R_{\vartheta_{k_1}^{(1)}})$$

and, since the relations $\overline{R_j}$ do not occur in φ_{1-j}, it follows that $(\mathfrak{A}, \overline{R}) \models_{X_j} \varphi_j^\star$ for $j \in \{0,1\}$. Therefore, $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$.

"$(i) \Longleftarrow (ii)$": For the other direction, let there be $\approx$-closed relations $\overline{R}$ such that $(\mathfrak{A}, \overline{R}) \models_X \varphi^\star$ and, hence, there are teams X_0, X_1 such that $X = X_0 \cup X_1$ and $(\mathfrak{A}, \overline{R}) \models_{X_j} \varphi_j^\star$ for $j \in \{0,1\}$. By induction hypothesis, this implies that $\mathfrak{A} \models_{X_j} \varphi_j$ for $j \in \{0,1\}$, whence it follows that $\mathfrak{A} \models_X \varphi$.

The case where $\varphi = \varphi_0 \wedge \varphi_1$ is similar to the previous one, and the cases where $\varphi = \forall x\psi$ or $\varphi = \exists x\psi$ are trivial. ◀

Proof of Lemma 9. Let $\mathfrak{A} \in (\tau, \approx)$ and X be some team over $\mathfrak{A}$ with $\mathrm{free}(\varphi) \subseteq \mathrm{dom}(X)$. Recall that $\vartheta_1, \dots, \vartheta_k$ are the inclusion atoms that occur in φ. Our goal is now to prove that the following statements are equivalent for every subformula ϑ of φ:

(1) $\mathfrak{A} \models_X \vartheta$

(2) There are $\approx$-closed relations $\overline{R} = (R_{\vartheta_1}, \dots, R_{\vartheta_k})$ over $\mathfrak{A}$ such that $(\mathfrak{A}, \overline{R}) \models_X \vartheta^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists some $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$.

Whenever we are proving that (2) holds, we can often use that $\vartheta^{(i)}$ and $\vartheta^\star$ are equivalent, if $\vartheta_i = \overline{x_i} \subseteq_\approx \overline{y_i}$ does not occur in ϑ. Hence, we only need to prove that $\mathfrak{A} \models_X \vartheta^\star$ holds and that for every $i \in \{1, \dots, k\}$ such that ϑ_i occurs in ϑ and every $\overline{a} \in R_{\vartheta_i}$ there exists some $s_{i,\overline{a}} \in X$ with $\mathfrak{A} \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}(\overline{z_i})$.

For the empty team $X = \varnothing$, there is nothing to prove, because $\mathfrak{A} \models_\varnothing \vartheta$ follows from the empty team property and the empty relations trivially satisfy the conditions stated in (2). From now on we only consider non-empty teams X in this proof, which proceeds now by induction:

Case $\vartheta = \overline{v} \subseteq_\approx \overline{w}$: Then there exists a unique $\ell \in \{1, \dots, k\}$ such that $\vartheta_\ell = \vartheta$ and $R_{\vartheta_\ell} = R_\vartheta$. Recall that we have defined $\vartheta^\star := R_\vartheta\overline{v}$ and

$$\vartheta^{(i)} := \begin{cases} \vartheta^\star, & \text{if } i \neq \ell \\ \vartheta^\star \wedge \overline{w} \approx \overline{z_i}, & \text{if } i = \ell. \end{cases}$$

"(1) $\Longrightarrow$ (2)": Suppose $\mathfrak{A} \models_X \overline{v} \subseteq_{\approx} \overline{w}$. Then $X(\overline{v}) \subseteq_{\approx} X(\overline{w})$ and, this is why, setting $R_\vartheta := X(\overline{w})_{\approx}$ leads to $(\mathfrak{A}, R_\vartheta) \models_X R_\vartheta \overline{v} = \vartheta^\star$. Moreover, for every $\overline{a} \in R_\vartheta = X(\overline{w})_{\approx}$ there exists, by definition of $X(\overline{w})_{\approx}$, an assignment $s \in X$ such that $s(\overline{w}) \approx \overline{a}$ and, hence, it holds that $(\mathfrak{A}, R_\vartheta) \models_{s[\overline{z_\ell} \mapsto \overline{a}]} \overline{w} \approx \overline{z_\ell}$, which leads to $(\mathfrak{A}, R_\vartheta) \models_{s[\overline{z_\ell} \mapsto \overline{a}]} \vartheta^{(\ell)}$. Since the other relations R_{ϑ_i} for $i \neq \ell$ occur neither in $\vartheta^\star$ nor $\vartheta^{(i)}$, it does not matter what values they contain.

"(1) $\Longleftarrow$ (2)": For the converse direction, we assume that there are $\approx$-closed relations $\overline{R}$ such that $(\mathfrak{A}, \overline{R}) \models_X R_{\vartheta_\ell} \overline{v}$ and that for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there is some $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$. In particular, for every $\overline{a} \in R_{\vartheta_\ell}$ holds

$$(\mathfrak{A}, \overline{R}) \models_{s_{\ell,\overline{a}}[\overline{z_\ell} \mapsto \overline{a}]} \vartheta^{(\ell)} = R_{\vartheta_\ell} \overline{v} \land \overline{w} \approx \overline{z_\ell}.$$

Our goal is to prove that $\mathfrak{A} \models_X \overline{v} \subseteq_{\approx} \overline{w}$. Towards this end, let $s \in X$. Because $(\mathfrak{A}, \overline{R}) \models_X R_{\vartheta_\ell} \overline{v}$, it follows that $s(\overline{v}) \in R_{\vartheta_\ell}$. So for $s' := s_{\ell, s(\overline{v})} \in X$ holds $(\mathfrak{A}, \overline{R}) \models_{s'[\overline{z_\ell} \mapsto s(\overline{v})]} \overline{w} \approx \overline{z_\ell}$. Therefore, $s'(\overline{w}) \approx s(\overline{v})$. Since $s \in X$ was chosen arbitrarily, this proves that $\mathfrak{A} \models_X \overline{v} \subseteq_{\approx} \overline{w}$.

Case ϑ is an FO-literal: Then we have $\vartheta^\star := \vartheta =: \vartheta^{(i)}$. For arbitrary (not necessarily $\approx$-closed) relations $\overline{R}$ holds

$$\begin{aligned} \mathfrak{A} \models_X \vartheta \underset{X \neq \varnothing}{\Longleftrightarrow} & (\mathfrak{A}, \overline{R}) \models_s \vartheta \text{ for every } s \in X \text{ and} \\ & \text{for every } i \in \{1, \dots, k\},\ \overline{a} \in R_{\vartheta_i} \text{ exists } s \in X \\ & \text{such that } (\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \vartheta = \vartheta^{(i)} \end{aligned}$$

Case $\vartheta = \psi_0 \lor \psi_1$: Let $\vartheta_1^{(j)}, \dots, \vartheta_{k_j}^{(j)}$ be the inclusion atoms that occur in ψ_j.

"(1) $\Longrightarrow$ (2)": First we assume that $\mathfrak{A} \models_X \vartheta$. Then there are two teams Y_0, Y_1 such that $X = Y_0 \cup Y_1$ and $\mathfrak{A} \models_{Y_j} \psi_j$ for $j \in \{0, 1\}$.

By induction hypothesis, there are tuples of $\approx$-closed relations $\overline{R_{(j)}} = (R^{(j)}_{\vartheta_1}, \dots, R^{(j)}_{\vartheta_k})$ such that $(\mathfrak{A}, \overline{R_{(j)}}) \models_{Y_j} \psi_j$ (for $j \in \{0,1\}$) and for every $i \in \{1, \dots, k\}$ and every $\overline{a} \in R^{(j)}_{\vartheta_i}$ there exists some $s \in Y_j$ such that $(\mathfrak{A}, \overline{R_{(j)}}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \psi_j^{(i)}$. Let $\overline{R} = (R_{\vartheta_1}, \dots, R_{\vartheta_k})$ be a tuple of $\approx$-closed relations such that $R_{\vartheta_i} = R^{(j)}_{\vartheta_i} \Longleftrightarrow \vartheta_i$ occurs in ψ_j.[3]

We are going to prove that $(\mathfrak{A}, \overline{R}) \models_X \vartheta^\star$ and that for every $i \in \{1, \dots, k\}$ such that ϑ_i occurs in ϑ and every $\overline{a} \in R_{\vartheta_i}$, there exists some $s \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$. Since $(\mathfrak{A}, \overline{R_{(j)}}) \models_{Y_j} \psi_j^\star$, we also have $(\mathfrak{A}, \overline{R}) \models_{Y_j} \psi_j^\star$, because whenever a relation symbol R_{ϑ_i} occurs in ψ_j, it must be the case that ϑ_i occurs in ψ_j and, hence, $R_{\vartheta_i} = R^{(j)}_{\vartheta_i}$. Additionally we still have $X = Y_0 \cup Y_1$ and, thus, $(\mathfrak{A}, \overline{R}) \models_X \vartheta^\star$.

Towards proving the second part, let $i \in \{1, \dots, k\}$ such that ϑ_i occurs in ϑ and $\overline{a} \in R_{\vartheta_i}$. There must be some (unique) $j \in \{0, 1\}$ such that ϑ_i occurs in ψ_j and $R_{\vartheta_i} = R^{(j)}_{\vartheta_i}$. We know already that there exists some $s \in Y_j$ that satisfies $(\mathfrak{A}, \overline{R^{(j)}}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \psi_j^{(i)}$, which implies that $(\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \psi_j^{(i)}$. Furthermore, we have that $\vartheta^{(i)} := \psi_j^{(i)}$ and, consequently, it follows that $(\mathfrak{A}, \overline{R}) \models_{s[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$, which is exactly what we wanted to achieve.

"(1) $\Longleftarrow$ (2)": Suppose that there are $\approx$-closed relations $\overline{R}$ such that $(\mathfrak{A}, \overline{R}) \models_X \vartheta^\star$ and that for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists some $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$. Then there are some teams Y_0, Y_1 such that $Y = Y_0 \cup Y_1$ and $(\mathfrak{A}, \overline{R}) \models_{Y_j} \psi_j^\star$ for $j \in \{0, 1\}$. Furthermore, by definition of $\vartheta^{(i)}$, for every $i \in \{1, \dots, k\}$ such that ϑ_i occurs in ϑ, there exists a (unique) $j(i) \in \{0, 1\}$ with $\vartheta^{(i)} = \psi^{(i)}_{j(i)}$. For $j \in \{0, 1\}$ let

$$Y'_j := \{s_{i,\overline{a}} \,:\, i \in \{1, \dots, k\},\ \vartheta_i \text{ occurs in } \vartheta,\ \overline{a} \in R_{\vartheta_i} \text{ and } j(i) = j\} \ (\subseteq X).$$

[3] Such $\overline{R}$ exists, because ϑ_i cannot occur in both ψ_0 and ψ_1.

It follows that $(\mathfrak{A}, \overline{R}) \models_s \psi_j^{(i)}$ for every $s \in Y_j'$, because every $s \in Y_j$ has the form $s = s_{i,\overline{a}}$ and we have that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}} \vartheta^{(i)} = \psi_{j(i)}^{(i)} (= \psi_j^{(i)})$. Since $\psi_j^{(i)} \models \psi_j^\star$, we can conclude that also $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi_j^\star$. This, together with the flatness property of FO, implies that $(\mathfrak{A}, \overline{R}) \models_{Z_j} \psi_j^\star$ where $Z_j := Y_j \cup Y_j'$.

For $j \in \{0, 1\}$ let $\overline{R_{(j)}} = (R_{\vartheta_1}^{(j)}, \dots, R_{\vartheta_k}^{(j)})$ be given by

$$R_{\vartheta_i}^{(j)} := \begin{cases} R_{\vartheta_i}, & \text{if } \vartheta_i \text{ is a subformula of } \psi_j \\ \varnothing, & \text{otherwise.} \end{cases}$$

Because the relation symbol R_{ϑ_i} occurs in $\psi_j^\star$ if, and only if ϑ_i is a subformula of ψ_j, we still have $(\mathfrak{A}, \overline{R_{(j)}}) \models_{Z_j} \psi_j^\star$ for $j \in \{0, 1\}$. Furthermore, for every $j \in \{0, 1\}$, every $i \in \{1, \dots, k\}$ and every $\overline{a} \in R_{\vartheta_i}^{(j)}$ it must be the case that ϑ_i is a subformula of ψ_j (otherwise we would have $R_{\vartheta_i}^{(j)} = \varnothing$, but this contradicts $\overline{a} \in R_{\vartheta_i}^{(j)}$) and, thus, it follows that $\vartheta^{(i)} = \psi_j^{(i)}$ and, therefore, $(\mathfrak{A}, \overline{R_{(j)}}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi_j^{(i)}$, because we have $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$ and $s_{i,\overline{a}} \in Y_j' \subseteq Z_j$.

This is the reason, why we are allowed to use the induction hypothesis, which yields us that $\mathfrak{A} \models_{Z_j} \psi_j$ for $j \in \{0, 1\}$. Consequently, it follows that $\mathfrak{A} \models_{Z_0 \cup Z_1} \vartheta$.

It is easy to observe that $Z_0 \cup Z_1 = Y_0 \cup Y_1 \cup Y_0' \cup Y_1' = X$, because $Y_0', Y_1' \subseteq X$ and $X = Y_0 \cup Y_1$. As a result, we obtain that $\mathfrak{A} \models_X \vartheta$.

Case $\vartheta = \psi_0 \wedge \psi_1$: Similar and even easier than the previous case!

Case $\vartheta = \exists x \psi$: Recall that we have defined $\vartheta^\star := \exists x \psi^\star$ and $\vartheta^{(i)} = \exists x \psi^{(i)}$ for every $i \in \{1, \dots, k\}$. We only prove "(1) $\Longleftarrow$ (2)", since the other direction is quite trivial.

Suppose that there are $\approx$-closed relations $\overline{R}$ such that $(\mathfrak{A}, \overline{R}) \models_X \exists x \psi^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists an $s_{i,\overline{a}} \in X$ such that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \exists x \psi^{(i)}$. Then there is a function $F : X \to \mathcal{P}(A) \setminus \{\varnothing\}$ such that for $Y := X[x \mapsto F]$ holds $(\mathfrak{A}, \overline{R}) \models_Y \psi^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_i$ there exists some $b_{i,\overline{a}} \in A$ such that $(\mathfrak{A}, \overline{R}) \models_{s'_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi^{(i)}$ where $s'_{i,\overline{a}} = s_{i,\overline{a}}[x \mapsto b_{i,\overline{a}}]$. Let $Y' := \{s'_{i,\overline{a}} : i \in \{1, \dots, k\}, \overline{a} \in R_{\vartheta_i}\}$. Due to $\psi^{(i)} \models \psi^\star$ and the flatness property of FO, it follows that $(\mathfrak{A}, \overline{R}) \models_Z \psi^\star$ for $Z := Y \cup Y'$. Furthermore, for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ we have $s'_{i,\overline{a}} \in Y' \subseteq Z$ with $(\mathfrak{A}, \overline{R}) \models_{s'_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi^{(i)}$. Thus, we can apply the induction hypothesis with Z and $\widetilde{\vartheta}$ to obtain that $\mathfrak{A} \models_Z \psi$. By definition of Z we have $Z \upharpoonright \mathrm{dom}(X) = (Y \upharpoonright \mathrm{dom}(X)) \cup (Y' \upharpoonright \mathrm{dom}(X))$. Furthermore, it is the case that $Y \upharpoonright \mathrm{dom}(X) = X[x \mapsto F] \upharpoonright \mathrm{dom}(X) = X$ (recall that we assume that no variable is quantified twice and that $\mathrm{dom}(X) = \mathrm{free}(\vartheta)$) and $Y' \upharpoonright \mathrm{dom}(X) \subseteq X$, because every $s' \in Y'$ has the form $s' = s'_{i,\overline{a}} = s_{i,\overline{a}}[x \mapsto b_{i,\overline{a}}]$ where $s_{i,\overline{a}} \in X$. Therefore, $Z \upharpoonright \mathrm{dom}(X) = X$ and, hence, it follows that $\mathfrak{A} \models_X \exists x \psi = \vartheta$.

Case $\vartheta = \forall x \psi$: Recall that we have defined $\vartheta^\star := \forall x \psi^\star$ and $\vartheta^{(i)} := \exists x(\psi^{(i)}) \wedge \forall x(\psi^\star)$.

"(1) $\Longrightarrow$ (2)": Let $\mathfrak{A} \models_X \forall x \psi$. Then $\mathfrak{A} \models_Y \psi$ where $Y := X[x \mapsto A]$. By induction hypothesis, there are $\approx$-closed relations $\overline{R}$ such that $(\mathfrak{A}, \overline{R}) \models_Y \psi^\star$ and for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists an $s'_{i,\overline{a}} \in Y$ that satisfies $(\mathfrak{A}, \overline{R}) \models_{s'_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi^{(i)}$.

Since $Y = X[x \mapsto A]$, it follows that $(\mathfrak{A}, \overline{R}) \models_X \forall x \psi^\star = \vartheta^\star$. We have already mentioned above that $\vartheta^\star \in \mathrm{FO}$. So we can use the flatness property, which leads to $(\mathfrak{A}, \overline{R}) \models_s \forall x \psi^\star$ for every $s \in X$. In particular, this holds for the assignments $s_{i,\overline{a}} := (s'_{i,\overline{a}} \upharpoonright \mathrm{dom}(X)) \in X$. This is why, we have $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}} \forall x \psi^\star$. Furthermore, from $(\mathfrak{A}, \overline{R}) \models_{s'_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \psi^{(i)}$ follows that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \exists x \psi^{(i)}$. Consequently, we can conclude that $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}} \exists x(\psi^{(i)}) \wedge \forall x(\psi^\star) = \vartheta^{(i)}$ for every $i \in \{1, \dots, k\}$, $\overline{a} \in R_{\vartheta_i}$ and we have that $(\mathfrak{A}, \overline{R}) \models_X \vartheta^\star$.

"(1) $\Longleftarrow$ (2)": Suppose that there are $\approx$-closed relations $\overline{R}$ satisfying $(\mathfrak{A}, \overline{R}) \models_X \forall x\psi^\star$ and for every $i \in \{1, \ldots, k\}$, $\overline{a} \in R_{\vartheta_i}$ there exists some $s_{i,\overline{a}} \in X$ such that

$$(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)} = \exists x(\psi^{(i)}) \wedge \vartheta^\star.$$

Let $Y := X[x \mapsto A]$. Then we have $(\mathfrak{A}, \overline{R}) \models_Y \psi^\star$ (because $(\mathfrak{A}, \overline{R}) \models_X \forall x\psi^\star$). Furthermore, for every $i \in \{1, \ldots, k\}$, $\overline{a} \in A$ there exists some $b_{i,\overline{a}} \in A$ such that $(\mathfrak{A}, \overline{R}) \models_{s'_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \vartheta^{(i)}$ where $s'_{i,\overline{a}} := s_{i,\overline{a}}[x \mapsto b_{i,\overline{a}}] \in Y$ (because $(\mathfrak{A}, \overline{R}) \models_{s_{i,\overline{a}}[\overline{z_i} \mapsto \overline{a}]} \exists x\psi^{(i)}$). So, by induction hypothesis, it follows that $\mathfrak{A} \models_Y \psi$ and, because of $Y = X[x \mapsto A]$, we obtain that $\mathfrak{A} \models_X \forall x\psi = \vartheta$. ◀

A.2 $\Sigma_1^1(\approx)$ on restricted classes of structures

Proof of Theorem 25. It suffices to prove this for $p = 2$. Let $\tau = \{E, \approx\}$. Consider the following problem: $\mathcal{C} := \{\mathfrak{A} \in \mathcal{K}_{\leq 2} : (A, E^{\mathfrak{A}})$ is not connected$\}$. By using the method of Ehrenfeucht-Fraïssé we will show that $\mathcal{C}$ is not definable in FO$(\subseteq_\approx, |_\approx)$.

For every $m > 3$ let $\mathfrak{A}_m := (A_m, E^{\mathfrak{A}_m}, \approx)$ and $\mathfrak{B}_m := (B_m, E^{\mathfrak{B}_m}, \approx)$ where $A_m := \{0, \ldots, m-1\} \cup \{0', \ldots, (m-1)'\} =: B_m$ and $E^{\mathfrak{A}_m} := E_+^{\mathfrak{A}_m} \cup E_-^{\mathfrak{A}_m}$ with

$$E_+^{\mathfrak{A}_m} := \{(i, j), (i', j') : j = i + 1 \ (\mathrm{mod}\ m)\}$$

and $E_-^{\mathfrak{A}_m} := \{(w, v) : (v, w) \in E_+^{\mathfrak{A}_m}\}$. Similary, $E^{\mathfrak{B}_m} := E_+^{\mathfrak{B}_m} \cup E_-^{\mathfrak{B}_m}$ where $E_+^{\mathfrak{B}_m} := \{(0, 1), (1, 2), \ldots, (m-2, m-1), (m-1, 0'), (0', 1'), \ldots, ((m-2)', (m-1)'), ((m-1)', 0)\}$ and $E_-^{\mathfrak{B}_m} := \{(w, v) : (v, w) \in E_+^{\mathfrak{B}_m}\}$. $\approx$ is in both structures defined such that $[i]_\approx = \{i, i'\}$ for every $i \in \{0, \ldots, m-1\}$. In other words, $\mathfrak{A}_m$ consists of two cycles $(0, 1, \ldots, m-1, 0)$ and $(0', 1', \ldots, (m-1)', 0')$ of length m, while $\mathfrak{B}_m$ is a single cycle $(0, 1, \ldots, m-1, 0', 1', \ldots, (m-1)', 0)$ of length $2m$.

For every $v \in \{0, 1, \ldots, m-1, 0', 1', \ldots, (m-1)'\}$ there are uniquely determined $s^{\mathfrak{A}_m}(v)$ and $s^{\mathfrak{B}_m}(v)$ such that $(v, s^{\mathfrak{A}_m}(v)) \in E_+^{\mathfrak{A}_m}$ and $(v, s^{\mathfrak{B}_m}(v)) \in E_+^{\mathfrak{B}_m}$. Similarly, there are exists uniquely determined predecessors $(s^{\mathfrak{A}_m})^{-1}(v)$ and $(s^{\mathfrak{B}_m})^{-1}(v)$ with $(v, (s^{\mathfrak{A}_m})^{-1}(v)) \in E_-^{\mathfrak{A}_m}$ and $(v, (s^{\mathfrak{A}_m})^{-1}(v)) \in E_-^{\mathfrak{B}_m}$. We define for every $v \in A_m$, every $w \in B_m$ and every $k \in \mathbb{Z}$

$$v +_{\mathfrak{A}_m} k := (s^{\mathfrak{A}_m})^k(v) \text{ and } w +_{\mathfrak{B}_m} k := (s^{\mathfrak{B}_m})^k(w).$$

We are going omit $\mathfrak{A}_m$ and $\mathfrak{B}_m$ as a subscript, when it is clear from the context that v belongs to $\mathfrak{A}_m$ resp. $\mathfrak{B}_m$.

For $v, w \in A_m$ we define $\mathrm{dist}_{\mathfrak{A}_m}(v, w)$ to be the minimal number $n \in \mathbb{N}$ such that $v + n = w$ or $v - n = w$, or ∞, if no such number $n \in \mathbb{N}$ exists. $\mathrm{dist}_{\mathfrak{B}_m}(v, w)$ is defined analogously. Please note, that $\mathrm{dist}_{\mathfrak{A}_m}(v, w) = \mathrm{dist}_{\mathfrak{A}_m}(w, v)$ and $\mathrm{dist}_{\mathfrak{B}_m}(v, w) = \mathrm{dist}_{\mathfrak{B}_m}(w, v)$. Furthermore, for every $a \in \{0, 1, \ldots, m-1, 0', 1', \ldots, (m-1)'\}$ and every $b, c \in \mathbb{Z}$ holds,

$$(a +_{\mathfrak{A}_m} b) +_{\mathfrak{A}_m} c = a +_{\mathfrak{A}_m} (b + c) \text{ and } (a +_{\mathfrak{B}_m} b) +_{\mathfrak{B}_m} c = a +_{\mathfrak{B}_m} (b + c).$$

It is easy to see that $\mathrm{dist}(v_1, v_3) \leq \mathrm{dist}(v_1, v_2) + \mathrm{dist}(v_2, v_3)$ for every v_1, v_2, v_3 from A_m or B_m. Furthermore, $v \approx w$ implies that $s^{\mathfrak{A}_m}(v) \approx s^{\mathfrak{B}_m}(v)$ and $(s^{\mathfrak{A}_m})^{-1}(v) \approx (s^{\mathfrak{B}_m})^{-1}(v)$. This observation leads to the following corollary.

▶ Claim 26. Let $v \in A_m, w \in B_m$ with $v \approx w$. Then $v + k \approx w + k$ for every $k \in \mathbb{Z}$.

For every $i, j, q \in \mathbb{N}$ we write $i \approx_q j$ if, and only if $i = j$ or $i \geq q \leq j$. Given two assignments $s : \{x_1, \ldots, x_\ell\} \to A_m$ and $t : \{x_1, \ldots, x_\ell\} \to B_m$, we write $s \approx_q t$ if, and only if $s(x_i) \approx t(x_i)$ (which is equivalent to: $s(x_i), t(x_i) \in \{n, n'\}$ for some $n \in \{0, \ldots, m-1\}$) and $\mathrm{dist}_{\mathfrak{A}_m}(s(x_i), s(x_j)) \approx_q \mathrm{dist}_{\mathfrak{B}_m}(t(x_i), t(x_j))$ holds for every $i, j \in \{1, \ldots, \ell\}$.

▶ **Lemma 27.** *Let $m > 2^{n+2}$ and $0 \leq \ell \leq k < n$. Furthermore, let $s : \{x_1, \dots, x_\ell\} \to A_m$ and $t : \{x_1, \dots, x_\ell\} \to B_m$ be two assignments with $s \approx_{2^{n+1-k}} t$. Then:*

(1) *For every $a \in A_m$ there exists some $b = b(s, t, a) \in B_m$ such that*

$$s' := s[x_{\ell+1} \mapsto a] \approx_{2^{n-k}} t[x_{\ell+1} \mapsto b] =: t'.$$

(2) *For every $b \in B_m$ there exists some $a = a(s, t, b) \in A_m$ such that*

$$s' := s[x_{\ell+1} \mapsto a] \approx_{2^{n-k}} t[x_{\ell+1} \mapsto b] =: t'.$$

Furthermore, for two teams X, Y over $\mathfrak{A}_m, \mathfrak{B}_m$ with $\mathrm{dom}(X) = \{x_1, \dots, x_\ell\} = \mathrm{dom}(Y)$ we write $X \approx_q Y$ if, and only if for every $s \in X$ there exists some $t \in Y$ and, conversely, for every $t \in Y$ there exists some $s \in X$ such that $s \approx_q t$.

▶ Claim 28. Let $n, m \in \mathbb{N}$ with $m > 2^{n+2}$. Duplicator has a winning strategy in $\mathcal{G}_n(\mathfrak{A}_m, \mathfrak{B}_m)$.

Thus we have $\mathfrak{A}_m \Rrightarrow_n \mathfrak{B}_m$ for every $m > 2^{n+2}$. Using very similar arguments, it is possible to prove that $\mathfrak{B}_m \Rrightarrow_n \mathfrak{A}_m$. Furthermore, we have $\mathfrak{A}_m \in \mathcal{C}$ and $\mathfrak{B}_m \notin \mathcal{C}$. This proves that $\mathcal{C}$ is not definable in $\mathrm{FO}(\subseteq_\approx, |_\approx)$ (because φ is unable to distinguish between $\mathfrak{A}_m$ and $\mathfrak{B}_m$ for every $m > 2^{\mathrm{depth}(\varphi)+2}$). On the other hand, $\mathcal{C}$ is definable in Σ_1^1 by the sentence $\exists X \exists x \exists y (Xx \land \neg Xy \land \forall u \forall v (Xu \land Euv \to Xv))$. This concludes the proof of $\mathrm{FO}(\subseteq_\approx, |_\approx) < \Sigma_1^1$. $\mathrm{FO} < \mathrm{FO}(\subseteq_\approx, |_\approx)$ follows from the fact that $\mathrm{FO}(|_\approx) \equiv \mathrm{FO}(\mathrm{dep}_\approx)$ and that the sentence

$$\forall x \exists y \forall x' \exists y' (\mathrm{dep}_\approx(x, y) \land \mathrm{dep}_\approx(x', y') \land x \not\approx y \land (x \not\approx x' \lor y \approx y') \land (x \not\approx y' \lor y \approx x'))$$

expresses that a even number of equivalence classes exists. Using standard Ehrenfeucht-Fraïssé arguments, it is not difficult to prove, that $\mathcal{C}$ is not FO-definable. ◀