On the Union Closed Fragment of Existential Second-Order Logic and Logics with Team Semantics

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

Richard Wilke
Mathematical Foundations of Computer Science, RWTH Aachen University, Aachen, Germany
wilke@logic.rwth-aachen.de

Abstract

We present syntactic characterisations for the union closed fragments of existential second-order logic and of logics with team semantics. Since union closure is a semantical and undecidable property, the normal form we introduce enables the handling and provides a better understanding of this fragment. We also introduce inclusion-exclusion games that turn out to be precisely the corresponding model-checking games. These games are not only interesting in their own right, but they also are a key factor towards building a bridge between the semantic and syntactic fragments. On the level of logics with team semantics we additionally present restrictions of inclusion-exclusion logic to capture the union closed fragment. Moreover, we define a team based atom that when adding it to first-order logic also precisely captures the union closed fragment of existential second-order logic which answers an open question by Galliani and Hella.

2012 ACM Subject Classification Theory of computation → Higher order logic

Keywords and phrases Higher order logic, Existential second-order logic, Team semantics, Closure properties, Union closure, Model-checking games, Syntactic charactisations of semantical fragments

Digital Object Identifier 10.4230/LIPIcs.CSL.2020.25

Related Version A full version of the paper is available at `https://arxiv.org/abs/1910.06057`.

Funding *Matthias Hoelzel*: This author is supported by the German Research Foundation (DFG).
Richard Wilke: This author is supported by the DFG, Research Training Group 2236 UnRAVeL.

1 Introduction

One branch of model theory engages with the characterisation of semantical fragments, which typically are undecidable, as syntactical fragments of the logics under consideration. Prominent examples are van Benthem's Theorem characterising the bisimulation invariant fragment of first-order logic as the modal-logic [10] or preservation theorems like the Łoś-Tarski Theorem, which states that formulae preserved in substructures are equivalent to universal formulae [6]. In this paper we consider formulae $\varphi(X)$ of existential second-order logic, Σ^1_1, in a free relational variable X and investigate the property of being closed under unions, meaning that whenever a family of relations X_i all satisfy φ, then their union $\bigcup_i X_i$ should also do so. Certainly closure under unions is an undecidable property. We provide a syntactical characterisation of all formulae of existential second-order logic obeying this property via a normal form called *myopic*-Σ^1_1, a notion based on ideas of Galliani and Hella [2]. By Fagin's Theorem, Σ^1_1 is the logical equivalent of the complexity class NP which highlights the importance to understand its fragments. Towards this end we employ game theoretic concepts and introduce a novel game type, called *inclusion-exclusion games*, suited for formulae $\varphi(X)$ with a free relational variable. In these games a strategy no longer is simply winning for one player – and hence proving whether a sentence is satisfied – but it is moreover adequate for a certain relation Y over $\mathfrak{A}$ showing that the formula is satisfied by $\mathfrak{A}$ and Y, in symbols $\mathfrak{A} \vDash \varphi(Y)$. We construct myopic-$\Sigma_1^1$ formulae that can define the winning regions of specifically those inclusion-exclusion games that are (semantically) closed under unions. Conceptually such games are eligible for any Σ_1^1-formula, but since our interest lies in those formulae that are closed under unions, we introduce a restricted version of such games, called *union games*, that precisely correspond to the model-checking games of union closed Σ_1^1-formulae. Consequently, the notion of union closure is captured on the level of formulae by the myopic fragment of Σ_1^1 and on the game theoretic level by union games.

Existential second-order logic has a tight connection to modern logics of dependence and independence that are based on the concept of teams, introduced by Hodges [7], and later refined by Väänänen in 2007 [9]. In contrast to classical logics, formulae of such a logic are evaluated against a set of assignments, called a *team*. One main characteristic of these logics is that dependencies between variables, such as "x depends solely on y", are expressed as atomic properties of teams. Widely used dependency atoms include dependence ($=\!(x, y)$), inclusion ($x \subseteq y$), exclusion ($x \mid y$) and independence ($x \bot y$). It is known that both independence logic FO($\bot$) and inclusion-exclusion logic FO($\subseteq, \mid$) have the same expressive power as full existential second-order logic Σ_1^1 [1]. The team in such logics corresponds to the free relational variable in existential second-order formulae, enabling us to ask the same questions about fragments with certain closure properties in both frameworks. One example of a well understood closure property is *downwards closure* stating that if a formula is satisfied by a team then it is also satisfied by all subteams (i.e. subsets of that team). It is well known that exclusion logic FO($\mid$) corresponds to the downwards closed fragment of Σ_1^1 [1, 8]. The issue of union closure is different. Galliani and Hella have shown that inclusion logic FO($\subseteq$) corresponds to greatest fixed-point-logic GFP$^+$ and, hence, by using the Immerman-Vardi Theorem, it captures all PTIME computable queries on ordered structures [2]. They also proved that every union closed dependency notion that itself is first-order definable (where the formula has access to a predicate for the team) is already definable in inclusion logic. However, there are union closed properties that are not definable in inclusion logic (think of a union closed NP property). For a concrete example we refer to the atom $\mathcal{R}$ from [2]. Thus Galliani and Hella asked the question whether there is a union closed atomic dependency notion β, such that the logic FO(β) captures precisely the union closed fragment of FO($\subseteq, \mid$). In the present work we answer this question positively with the aid of inclusion-exclusion games. Furthermore, we present a syntactical restriction of all FO($\subseteq, \mid$) formulae that also precisely describe the union closed fragment. This syntactical fragment corresponds to myopic-Σ_1^1 and is in harmony with the game theoretical view, which is described by union games.

Sections 3, 4 and 5 deal with second-order logic, while the other sections, 6 and 7, address logics with team semantics. In section three the central notion of this paper, inclusion-exclusion games, are introduced, which are used in section four to characterise the union closed fragment within existential second-order logic. Section five provides a restriction of the games specifically suited for this fragment. The sections dealing with team semantics can be read mostly independently of each other. Based on section four, section six describes the union closed fragment of inclusion-exclusion logic in terms of syntactical restrictions. The question of Galliani and Hella, whether there is a union closed atom that constitutes the union closed fragment, is answered positively in section seven, for which the reader should be familiar with union games introduced in section five.
Omitted proofs can be found in the full version or can be done without much effort.

2 Preliminaries

We assume familiarity with first-order logic and existential second-order logic, FO and Σ_1^1 for short. For a background we refer to the textbook [4].

The neighbourhood of a vertex v in a graph G is denoted by $\mathrm{N}_G(v)$. For a given τ-structure $\mathfrak{A}$ and formula $\varphi(\bar{x})$ we define $\varphi^{\mathfrak{A}} := \{\bar{a} : \mathfrak{A} \vDash \varphi(\bar{a})\}$, $\mathrm{free}(\varphi)$ is the set of free first-order variables and $\mathrm{subf}(\psi)$ is the set of subformulae of ψ. Notations like $\bar{v}, \bar{w}$ always indicate that $\bar{v} = (v_1, \ldots, v_k)$ and $\bar{w} = (w_1, \ldots, w_\ell)$ are some (finite) tuples. Here $k = |\bar{v}|$ and $\ell = |\bar{w}|$, so $\bar{v}$ is a k-tuple while $\bar{w}$ is an ℓ-tuple. We write $\{\bar{v}\}$ or $\{\bar{v}, \bar{w}\}$ as abbreviations for $\{v_1, \ldots, v_k\}$ resp. $\{v_1, \ldots, v_k, w_1, \ldots, w_\ell\}$ while $\{(\bar{v}), (\bar{w})\}$ is the set consisting of the two tuples $\bar{v}$ and $\bar{w}$ (as elements). The *concatenation* of $\bar{v}$ and $\bar{w}$ is $(\bar{v}, \bar{w}) := (v_1, \ldots, v_k, w_1, \ldots, w_\ell)$. The power set of a set A is denoted by $\mathcal{P}(A)$ and $\mathcal{P}^+(A) := \mathcal{P}(A) \setminus \{\varnothing\}$.

Team Semantics. A *team* X over $\mathfrak{A}$ is a *set* of assignments mapping a common domain $\mathrm{dom}(X) = \{\bar{x}\}$ of variables into A.[1] The restriction of X to some first-order formula $\varphi(\bar{x})$ is $X\restriction_\varphi := \{s \in X : \mathfrak{A} \vDash_s \varphi\}$. For a given subtuple $\bar{y} = (y_1, \ldots, y_\ell) \subseteq \bar{x}$ and every $s \in X$ we define $s(\bar{y}) := (s(y_1), \ldots, s(y_\ell))$. Furthermore, we frequently use $X(\bar{y}) := \{s(\bar{y}) : s \in X\}$, which is an ℓ-ary relation over $\mathfrak{A}$. For an assignment s, a variable x and $a \in A$ we use $s[x \mapsto a]$ to denote the assignment resulting from s by adding x to its domain (if it is not already contained) and declaring a as the image of x.

▶ **Definition 1.** *Let $\mathfrak{A}$ be a τ-structure, X a team of $\mathfrak{A}$. In the following λ denotes a first-order τ-literal and φ, ψ arbitrary formulae in negation normal form.*
- $\mathfrak{A} \vDash_X \lambda :\Longleftrightarrow \mathfrak{A} \vDash_s \lambda$ *for all* $s \in X$
- $\mathfrak{A} \vDash_X \varphi \wedge \psi :\Longleftrightarrow \mathfrak{A} \vDash_X \varphi$ *and* $\mathfrak{A} \vDash_X \psi$
- $\mathfrak{A} \vDash_X \varphi \vee \psi :\Longleftrightarrow \mathfrak{A} \vDash_Y \varphi$ *and* $\mathfrak{A} \vDash_Z \psi$ *for some* $Y, Z \subseteq X$ *such that* $Y \cup Z = X$
- $\mathfrak{A} \vDash_X \forall x\varphi :\Longleftrightarrow \mathfrak{A} \vDash_{X[x\mapsto A]} \varphi$
- $\mathfrak{A} \vDash_X \exists x\varphi :\Longleftrightarrow \mathfrak{A} \vDash_{X[x\mapsto F]} \varphi$ *for some* $F\colon X \to \mathcal{P}^+(A)$

Here $X[x \mapsto A] := \{s[x \mapsto a] : s \in X, a \in A\}$ *and* $X[x \mapsto F] := \{s[x \mapsto a] : s \in X, a \in F(s)\}$.

Team semantics for a first-order formula φ (without any dependency concepts) boils down to evaluating φ against every single assignment, i.e. more formally we have $\mathfrak{A} \vDash_X \varphi \iff \mathfrak{A} \vDash_s \varphi$ for *every* $s \in X$ (in usual Tarski semantics). This is also known as the flatness property of FO. The reason for considering teams instead of single assignments is that they allow the formalisation of dependency statements in the form of dependency atoms. Among the most common atoms are the following.

- $\mathfrak{A} \vDash_X =\!(\bar{x}, y) :\Longleftrightarrow s(\bar{x}) = s'(\bar{x})$ implies $s(y) = s'(y)$ for all $s, s' \in X$
- $\mathfrak{A} \vDash_X \bar{x} \subseteq \bar{y} :\Longleftrightarrow X(\bar{x}) \subseteq X(\bar{y})$
- $\mathfrak{A} \vDash_X \bar{x} \mid \bar{y} :\Longleftrightarrow X(\bar{x}) \cap X(\bar{y}) = \varnothing$
- $\mathfrak{A} \vDash_X \bar{x} \perp \bar{y} :\Longleftrightarrow$ for all $s, s' \in X$ exists $s'' \in X$ s.t. $s(\bar{x}) = s''(\bar{x})$ and $s'(\bar{y}) = s''(\bar{y})$

These are called *dependence* [9], *inclusion, exclusion* [1] and *independence* [5] atoms, respectively. When we speak about a logic that may use certain atomic dependency notions, for example inclusion, we denote it by writing $\mathrm{FO}(\subseteq)$ and so forth. These logics have the empty team property, which means that $\mathfrak{A} \vDash_\varnothing \varphi$ is always true. This is also the reason why *sentences* are not evaluated against $\varnothing$ but rather against $\{\varnothing\}$, which is the team consisting

[1] We use the corresponding Latin letters to denote universes of structures.

of the empty assignment. Let φ be a first-order formula and ψ be any formula of a logic with team semantics. We define $\varphi \to \psi$ as $\mathrm{nnf}(\neg\varphi) \vee (\varphi \wedge \psi)$ where $\mathrm{nnf}(\neg\varphi)$ is the negation normal form of $\neg\varphi$. It is easy to see that $\mathfrak{A} \models_X \varphi \to \psi \iff \mathfrak{A} \models_{X\restriction_\varphi} \psi$.

Union Closure. A formula φ of a logic with team semantics is said to be *union closed* if $\mathfrak{A} \models_{X_i} \varphi$ for all $i \in I$ implies $\mathfrak{A} \models_X \varphi$, where $X = \bigcup_{i\in I} X_i$. Analogously, a formula $\varphi(X)$ of Σ^1_1 with a free relational variable X is union closed if $\mathfrak{A} \models \varphi(X_i)$ for all i implies $\mathfrak{A} \models \varphi(X)$.

FO Interpretations. A first-order interpretation from σ to τ of arity k is a sequence $\mathcal{I} = (\delta, \varepsilon, (\psi_S)_{S\in\tau})$ of $\mathrm{FO}(\sigma)$-formulae, called the domain, equality and relation formulae respectively. We say that $\mathcal{I}$ interprets a τ-structure $\mathfrak{B}$ in some σ-structure $\mathfrak{A}$ and write $\mathfrak{B} \cong \mathcal{I}(\mathfrak{A})$ if and only if there exists a surjective function h, called the coordinate map, that maps $\delta^{\mathfrak{A}} = \{\bar{a} \in A^k : \mathfrak{A} \models \delta(\bar{a})\}$ to B preserving and reflecting the equalities and relations provided by ε and ψ_S, such that h induces an isomorphism between the quotient structure $(\delta^{\mathfrak{A}}, (\psi_S^{\mathfrak{A}})_{S\in\tau})/\varepsilon^{\mathfrak{A}}$ and $\mathfrak{B}$. A more detailed explanation can be found in [4]. For a τ-formula φ we associate the σ-formula $\varphi^{\mathcal{I}}$ by relativising quantifiers to δ, using ε as equality and ψ_S instead of S. We extend this translation to Σ^1_1 by the following rules for additional free/quantified relation symbols S.

- $(\exists S\vartheta)^{\mathcal{I}} \coloneqq \exists S^\star \big(\forall \bar{x}_1 \cdots \bar{x}_{\mathrm{ar}(S)} \big(S^\star \bar{x}_1 \cdots \bar{x}_{\mathrm{ar}(S)} \to \bigwedge_{j=1}^{\mathrm{ar}(S)} \delta(\bar{x}_j)\big) \wedge \vartheta^{\mathcal{I}}\big)$,
- $(Sv_1 \cdots v_{\mathrm{ar}(S)})^{\mathcal{I}} \coloneqq \exists \bar{w}_1 \cdots \bar{w}_{\mathrm{ar}(S)} \big(\bigwedge_{j=1}^{\mathrm{ar}(S)} (\delta(\bar{w}_j) \wedge \varepsilon(\bar{v}_j, \bar{w}_j)) \wedge S^\star \bar{w}_1 \cdots \bar{w}_{\mathrm{ar}(S)}\big)$.

An assignment $s\colon \{\bar{x}_1, \dots, \bar{x}_m\} \to A$ is well-formed (w.r.t. $\mathcal{I}$), if $s(\bar{x}_i) \in \delta^{\mathfrak{A}} (= \mathrm{dom}(h))$ for every $i = 1, \dots, m$. Such an assignment encodes $h \circ s\colon \{x_1, \dots, x_m\} \to B$ with $(h \circ s)(x_i) \coloneqq h(s(\bar{x}_i))$ which is an assignment over $\mathfrak{B}$. Similarly, a relation Q is well-formed (w.r.t. $\mathcal{I}$), if $Q \subseteq (\delta^{\mathfrak{A}})^\ell$ where $\ell = \frac{\mathrm{ar}(Q)}{k} \in \mathbb{N}$, and we define $h(Q) \coloneqq \{(h(\bar{a}_1), \dots, h(\bar{a}_\ell)) : (\bar{a}_1, \dots, \bar{a}_\ell) \in Q\}$, which is the ℓ-ary relation over $\mathfrak{B}$ that was described by Q. The connection between $\varphi^{\mathcal{I}}$ and φ is made precise in the well-known interpretation lemma.

▶ **Lemma 2** (Interpretation Lemma for Σ^1_1)**.** *Let $\varphi(S_1, \dots, S_n) \in \Sigma^1_1$. Let $R_i^\star \subseteq A^{k\cdot\mathrm{ar}(S_i)}$ for all i and $s\colon \{\bar{x}_1, \dots, \bar{x}_m\} \to A$ be well-formed. Then: $(\mathfrak{A}, R_1^\star, \dots, R_n^\star) \models_s \varphi^{\mathcal{I}} \iff (\mathfrak{B}, h(R_1^\star), \dots, h(R_n^\star)) \models_{h\circ s} \varphi$.*

3 Inclusion-Exclusion Games

Classical model-checking games are designed to express satisfiability of sentences, i.e. formulae without free variables. Since our focus lies on formulae in a free relational variable we are in need for a game that is able to not only express that a formula is satisfied, but moreover that it is satisfied by a certain relation. In the games we are about to describe a set of designated positions is present – called the *target set* – which corresponds to the full relation A^k (where the free relational variable has arity k). A winning strategy is said to be *adequate* for a subset X of the target positions, if the target vertices visited by it are X. On the level of logics this matches the relation satisfying the corresponding formula, i.e. there is a winning strategy adequate for X if and only if the formula is satisfied by X.

An *inclusion-exclusion game* $\mathcal{G} = (V, V_0, V_1, E, I, T, E_{\mathsf{ex}})$ is played by two players 0 and 1 where

- V_σ is the set of vertices of player σ,
- $V = V_0 \mathbin{\dot\cup} V_1$,
- $E \subseteq V \times V$ is a set of possible moves,
- $I \subseteq V$ is the (possibly empty) set of initial positions,

- $T \subseteq V$ is the set of target vertices and
- $E_{\mathsf{ex}} \subseteq V \times V$ is the exclusion condition, which defines the winning condition for player 0.[2]

The edges going into T, that is $E_{\mathsf{in}} := E \cap (V \times T)$, are called *inclusion edges*, while E_{ex} is the set of *exclusion edges* (sometimes also called conflicting pairs). Inclusion-exclusion games are *second-order* games, so instead of single plays we are more interested in sets of plays that are admitted by some winning strategy for player 0.

For a subset $X \subseteq T$ the aim of player 0 is to provide a winning strategy (which can be viewed as a set of plays respecting the exclusion condition and containing all possible strategies of player 1) such that the vertices of T that are visited by this strategy correspond precisely to X.

▶ **Definition 3.** *A* winning strategy *(for player 0) $\mathcal{S}$ is a possibly empty subgraph $\mathcal{S} = (W, F)$ of $G = (V, E)$ ensuring the following four consistency conditions.*

(i) *For every $v \in W \cap V_0$ holds $\mathrm{N}_{\mathcal{S}}(v) \neq \varnothing$.*
(ii) *For every $v \in W \cap V_1$ holds $\mathrm{N}_{\mathcal{S}}(v) = \mathrm{N}_G(v)$.*
(iii) *$I \subseteq W$.*
(iv) *$(W \times W) \cap E_{\mathsf{ex}} = \varnothing$.*

Intuitively, the conditions (i) and (ii) state that the strategy must provide at least one move from each node of player 0 used by the strategy but does not make assumptions about the moves that player 1 may make whenever the strategy plays a node belonging to player 1. In particular, the strategy must not play any terminal vertices that are in V_0. Furthermore, (iii) enforces that at least the initial vertices are contained while (iv) disallows playing with conflicting pairs $(v, w) \in E_{\mathsf{ex}}$, i.e. v and w must not coexist in any winning strategy for player 0. If $I = \varnothing$, then $(\varnothing, \varnothing)$ is the trivial winning strategy. Since we do *not* have a notion for a winning strategy for player 1, inclusion-exclusion games can be viewed as solitaire games.

Of course, the winning condition of an inclusion-exclusion game $\mathcal{G}$ is first-order definable. The formula $\varphi_{\mathrm{win}}(W, F)$ has the property that $\mathcal{G} \vDash \varphi_{\mathrm{win}}(W, F)$ if and only if (W, F) is a winning strategy for player 0 in $\mathcal{G}$, where

$$\begin{aligned}\varphi_{\mathrm{win}}(W, F) \;:=\; &\forall v(Wv \to ((V_0 v \wedge \exists w(Evw \wedge Ww \wedge Fvw)) \vee \\ &\qquad\qquad (V_1 v \wedge \forall w(Evw \to Ww \wedge Fvw)))) \wedge \\ &\qquad \forall v(Iv \to Wv) \wedge \forall v \forall w((Wv \wedge Ww) \to \neg E_{\mathsf{ex}} vw)\end{aligned}$$

describes the winning condition imposed on the graph (W, F).

We are mainly interested in the subset of target vertices that are visited by a winning strategy $\mathcal{S} = (W, F)$. More formally, $\mathcal{S}$ induces $\mathcal{T}(\mathcal{S}) := W \cap T$, which we also call the *target* of $\mathcal{S}$. This allows us to associate with every inclusion-exclusion game $\mathcal{G}$ the set of targets of winning strategies: $\mathcal{T}(\mathcal{G}) := \{\mathcal{T}(\mathcal{S}) : \mathcal{S}$ is a winning strategy for player 0 in $\mathcal{G}\}$.

Intuitively, as already pointed out, games of this kind will be the model-checking games for Σ^1_1-formulae $\varphi(X)$ that have a free relational variable X. Given a structure $\mathfrak{A}$ and such a formula, we are interested in the possible relations Y that satisfy the formula, in symbols $(\mathfrak{A}, Y) \vDash \varphi(X)$. We will construct the game such that Y satisfies φ if and only if there is a strategy of player 0 winning for the set $Y \subseteq T$, thus $\mathcal{T}(\mathcal{G}) = \{Y : (\mathfrak{A}, Y) \vDash \varphi\}$. It will be more convenient for our purposes that the target vertices of an inclusion-exclusion game are not required to be terminal positions. However it would be no restriction as it is easy to transform any given game into one that agrees on the (possible) targets, in which all target vertices are terminal.

[2] E_{ex} can always be replaced by the symmetric closure of E_{ex} without altering its semantics.

One can reduce the satisfiability problem of propositional logic to deciding whether player 0 has a winning strategy in an inclusion-exclusion game.

▶ **Theorem 4.** *The problem of deciding whether $X \in \mathcal{T}(\mathcal{G})$ for a finite inclusion-exclusion game $\mathcal{G}$ is* NPcomplete.

3.1 Model-Checking Games for Existential Second-Order Logic

In this section we define model-checking games for formulae $\varphi(X) \in \Sigma_1^1$ with a free relational variable. These games are inclusion-exclusion games whose target sets are precisely the sets of relations that satisfy $\varphi(X)$.

▶ **Definition 5.** *Let $\mathfrak{A}$ be a τ-structure and $\varphi(X) = \exists \bar{R}\varphi'(X, \bar{R}) \in \Sigma_1^1$ (in negation-normal form) where $\varphi'(X, \bar{R}) \in \mathrm{FO}(\tau \cup \{X, \bar{R}\})$ using a free relation symbol X of arity $r := \mathrm{ar}(X)$. The game $\mathcal{G}_X(\mathfrak{A}, \varphi) := (V, V_0, V_1, E, I, T, E_{\mathsf{ex}})$ consists of the following components:*

- $V := \{(\vartheta, s) : \vartheta \in \mathrm{subf}(\varphi'), s\colon \mathrm{free}(\vartheta) \to A\} \cup A^r$, $I := \{(\varphi', \varnothing)\}$, $T := A^r$,
- $V_1 := \{(\vartheta, s) : \vartheta = \forall y\gamma$ *or* $\vartheta = \gamma_1 \wedge \gamma_2\} \cup \{(\gamma, s) : \gamma$ *is a τ-literal and* $\mathfrak{A} \models_s \gamma\} \cup$
  $\{(\gamma, s) : \gamma = \neg X\bar{x}$ *or γ is a $\{\bar{R}\}$-literal*$\} \cup T$, $V_0 := V \setminus V_1$,
- $E := \{((\gamma \circ \vartheta, s), (\delta, s\restriction_{\mathrm{free}(\delta)})) : \circ \in \{\wedge, \vee\}, \delta \in \{\gamma, \vartheta\}\} \cup$
  $\{((X\bar{x}, s), s(\bar{x})) : X\bar{x} \in \mathrm{subf}(\varphi')\} \cup$
  $\{((Qx\gamma, s), (\gamma, s')) : Q \in \{\exists, \forall\}, s' = s[x \mapsto a], a \in A\}$,
- $E_{\mathsf{ex}} := \{((R_i\bar{x}, s), (\neg R_i\bar{y}, s')) : s(\bar{x}) = s'(\bar{y})\} \cup \{((\neg X\bar{x}, s), \bar{a}) : s(\bar{x}) = \bar{a}\}$.

These games capture the behaviour of existential second-order formulae which provides us with the following theorem.

▶ **Theorem 6.** $(\mathfrak{A}, X) \models \varphi(X) \iff$ *Player 0 has a winning strategy $\mathcal{S}$ in $\mathcal{G}_X(\mathfrak{A}, \varphi)$ with $\mathcal{T}(\mathcal{S}) = X$. Or, in other words: $\mathcal{T}(\mathcal{G}_X(\mathfrak{A}, \varphi)) = \{X \subseteq A^r : (\mathfrak{A}, X) \models \varphi(X)\}$.*

4 Characterising the Union Closed Formulae within Existential Second-Order Logic

In this section we investigate formulae $\varphi(X)$ of existential second-order logic that are closed under unions with respect to their free relational variable X. Union closure, being a semantical property of formulae, is undecidable. However, we present a *syntactical* characterisation of all such formulae via the following normal form.

▶ **Definition 7.** *A formula $\varphi(X) \in \Sigma_1^1$ is called* myopic *if $\varphi(X) = \forall\bar{x}(X\bar{x} \to \exists\bar{R}\varphi'(X, \bar{R}))$, where $\varphi' \in \mathrm{FO}$ and X occurs only positively*[3] *in φ'.*

Variants of myopic formulae have already been considered for first-order logic [2, Definition 19] and for greatest fixed-point logics [3, Theorem 24 and Theorem 26], but to our knowledge myopic Σ_1^1-formulae have not been studied so far.

Let $\mathcal{U}$ denote the set of all union closed Σ_1^1-formulae. To establish the claim that myopic formuale are a normal form of $\mathcal{U}$ we need to show that all myopic formulae are indeed closed under unions and, more importantly, that every union closed formula can be translated into an equivalent myopic formula. This translation is in particular constructive.

[3] That is under an even number of negations.

▶ **Theorem 8.** *$\varphi(X) \in \Sigma^1_1$ is union closed if and only if $\varphi(X)$ is equivalent to some myopic Σ^1_1-formula.*

We split the proof into two parts, the direction from right to left is handled in Proposition 9 and from left to right in Theorem 11.

▶ **Proposition 9.** *Every myopic formula is union closed.*

Proof. Let $\varphi = \forall \bar{x}(X\bar{x} \to \exists \bar{R}\varphi'(X, \bar{R}))$ and $(\mathfrak{A}, X_i) \vDash \varphi$ for all $i \in I$. We claim that $(\mathfrak{A}, X) \vDash \varphi$ for $X = \bigcup_{i \in I} X_i$. Let $\bar{a} \in X_i \subseteq X$. By assumption $(\mathfrak{A}, X_i) \vDash_{\bar{x} \mapsto \bar{a}} \exists \bar{R}\varphi'(X, \bar{R})$. A fortiori ($X$ occurs only positively in φ'), we obtain $(\mathfrak{A}, X) \vDash_{\bar{x} \mapsto \bar{a}} \exists \bar{R}\varphi'(X, \bar{R})$. Since $\bar{a}$ was chosen arbitrarily, this property holds for all $\bar{a} \in X$, hence the claim follows. ◀

For a fixed formula $\varphi(X)$ the corresponding game $\mathcal{G}_X$ can be constructed by a first-order interpretation depending of course on the current structure.

▶ **Lemma 10.** *Let $\varphi(X) = \exists \bar{R}\varphi'(X, \bar{R}) \in \Sigma^1_1$ where $\varphi' \in \mathrm{FO}(\tau \cup \{X, \bar{R}\})$ and $r := \mathrm{ar}(X)$. Then there exists a quantifier-free interpretation $\mathcal{I}$ such that $\mathcal{G}_X(\mathfrak{A}, \varphi) \cong \mathcal{I}(\mathfrak{A})$ for every structure $\mathfrak{A}$ (with at least two elements).*

▶ **Theorem 11.** *For every union closed formula $\varphi(X) \in \Sigma^1_1$ there is an equivalent myopic formula $\mu(X) \in \Sigma^1_1$.*

Proof. Let $\varphi(X) = \exists \bar{R}\varphi'(X, \bar{R}) \in \Sigma^1_1(\tau)$ be closed under unions, $\mathfrak{A}$ be a τ-structure and $\mathcal{G} := \mathcal{G}_X(\mathfrak{A}, \varphi)$ be the corresponding game. W.l.o.g. $\mathfrak{A}$ has at least two elements. By Theorem 6, we have that $\mathcal{T}(\mathcal{G}) = \{Y \subseteq A^r : \mathfrak{A} \vDash \varphi(Y)\}$ where $r := \mathrm{ar}(X)$. Since $\varphi(X)$ is union closed, it follows that $\mathcal{T}(\mathcal{G})$ is closed under unions as well. Now we observe that $\mathcal{T}(\mathcal{G})$ can be defined in the game $\mathcal{G}$ by the following myopic formula:

$$\begin{aligned} \varphi_{\mathcal{T}}(X) &:= \forall x(Xx \to \psi_{\mathcal{T}}(X, x)) \text{ where} \\ \psi_{\mathcal{T}}(X, x) &:= \exists W \exists F(\varphi_{\mathrm{win}}(W, F) \land Wx \land \forall y(Wy \land Ty \to Xy)) \end{aligned}$$

Here φ_{win} is the first-order formula verifying winning strategies. Please note that $\varphi_{\mathcal{T}}$ is indeed a myopic formula, since X occurs only positively in $\psi_{\mathcal{T}}$.

▷ Claim 12. For every $X \subseteq A^r$, $(\mathcal{G}, X) \vDash \varphi_{\mathcal{T}}(X) \iff X \in \mathcal{T}(\mathcal{G})$.

Proof. Assume that $(\mathcal{G}, X) \vDash \varphi_{\mathcal{T}}(X)$. By construction of $\varphi_{\mathcal{T}}$, for every $\bar{a} \in X$ there exists a winning strategy $\mathcal{S}_{\bar{a}} = (W_{\bar{a}}, F_{\bar{a}})$ with $\bar{a} \in W_{\bar{a}}$ and $\mathcal{T}(\mathcal{S}_{\bar{a}}) = W_{\bar{a}} \cap T \subseteq X$. It follows that $X = \bigcup_{\bar{a} \in X} \mathcal{T}(\mathcal{S}_{\bar{a}})$. Since $\mathcal{T}(\mathcal{G})$ is closed under unions, we also obtain that $X \in \mathcal{T}(\mathcal{G})$.

We want to remark that at this point the semantical property is translated into a syntactical one, as the formula only describes the correct winning strategy because the initial formula was closed under unions.

To conclude the proof of Claim 12, assume that $X \in \mathcal{T}(\mathcal{G})$. Then there exists a winning strategy $\mathcal{S} = (W, F)$ for player 0 with $\mathcal{T}(\mathcal{S}) = X$. Thus, for the quantifiers $\exists W \exists F$ we can (for all $\bar{a} \in X$) choose $\mathcal{S}$, which, obviously, satisfies the formula. ◁

There is a first-order interpretation $\mathcal{I}$ (of arity $n+m$) with $\mathcal{I}(\mathfrak{A}) \cong \mathcal{G}$ for some coordinate map $h\colon \delta^{\mathfrak{A}} \to V(\mathcal{G})$ and for every $\bar{a} \in T(\mathcal{G})$, $h^{-1}(\bar{a}) = \{(\bar{u}, \bar{a}, \bar{b}) \in A^{n+m} : \mathfrak{A} \vDash e_T(\bar{u}), \bar{b} \in A^{m-r}\}$ where $e_T(x_1, \ldots, x_n)$ is some quantifier-free first-order formula (for more details, we refer to the full version). By the interpretation lemma for Σ^1_1 (Lemma 2), for every $X \subseteq T(\mathcal{G})$,

$$(\mathfrak{A}, X^\star) \vDash \varphi^{\mathcal{I}}_{\mathcal{T}}(X^\star) \iff (\mathcal{G}, X) \vDash \varphi_{\mathcal{T}}(X) \qquad (1)$$

where $X^\star := h^{-1}(X)$ is a relation of arity $(n+m)$. Recall that every variable x occurring in $\varphi_\mathcal{T}$ is replaced by a tuple $\bar{x}$ of length $(n+m)$. Let $\bar{x} = (\bar{u}, \bar{v}, \bar{w})$ where $|\bar{u}| = n$, $|\bar{v}| = r$ and $|\bar{w}| = m - r$ and let

$$\mu(X) := \forall \bar{v}(X\bar{v} \rightarrow \forall \bar{u} \forall \bar{w}(e_T(\bar{u}) \rightarrow \psi^\star(X, \bar{u}, \bar{v}, \bar{w})))$$

where $\psi^\star$ is the formula that results from $\psi^\mathcal{I}_\mathcal{T}$ by replacing every occurrence of $X^\star \bar{u}'\bar{v}'\bar{w}'$ (where $|\bar{u}'| = n$, $|\bar{v}'| = r$ and $|\bar{w}'| = m - r$) by the formula $e_T(\bar{u}') \wedge X\bar{v}'$. By construction, this is a myopic formula, because X occurred only positively in $\psi^\mathcal{I}$ and, hence, $X^\star$ (resp. X) occurs only positively in $\psi^\mathcal{I}_\mathcal{T}$ (resp. $\psi^\star$).

Recall that, in the game $\mathcal{G} \cong \mathcal{I}(\mathfrak{A})$, every $X \subseteq T(\mathcal{G})$ is a unary relation over $\mathcal{G}$, while the elements of $T(\mathcal{G})$ themselves are r-tuples of A. Furthermore, we have that $h^{-1}(X) := \{(\bar{a}, \bar{b}, \bar{c}) \in A^n \times A^r \times A^{m-r} : \mathfrak{A} \models e_T(\bar{a}) \text{ and } \bar{b} \in X\}$. Because of this and $X^\star = h^{-1}(X)$, it follows that for every $s\colon \{\bar{u}', \bar{v}', \bar{w}'\} \rightarrow A$ holds

$$(\mathfrak{A}, X^\star) \models_s X^\star \bar{u}'\bar{v}'\bar{w}' \iff \mathfrak{A} \models e_T(s(\bar{u}')) \text{ and } s(\bar{v}') \in X \iff (\mathfrak{A}, X) \models_s e_T(\bar{u}') \wedge X\bar{v}'. \quad (2)$$

By construction of $\psi^\star$, these are the only subformulae in which $\psi^\mathcal{I}_\mathcal{T}$ and $\psi^\star$ differ from each other. As a result, the following claim is true:

▷ Claim 13. For every $X \subseteq A^r$ and every assignment $s\colon \mathrm{free}(\psi^\mathcal{I}_\mathcal{T}) \rightarrow A$, holds

$$(\mathfrak{A}, X^\star) \models_s \psi^\mathcal{I}_\mathcal{T}(X^\star, \bar{x}) \iff (\mathfrak{A}, X) \models_s \psi^\star(X, \bar{x}).$$

Recall that $\bar{x} = (\bar{u}, \bar{v}, \bar{w})$ where $|\bar{u}| = n, |\bar{v}| = r$ and $|\bar{w}| = (m - r)$. Now we can see that

$$\begin{aligned}
&(\mathfrak{A}, X^\star) \models \varphi^\mathcal{I}_\mathcal{T} = \forall \bar{x}(X^\star \bar{x} \rightarrow \psi^\mathcal{I}_\mathcal{T}(X^\star, \bar{x}))\\
\iff &(\mathfrak{A}, X^\star) \models_s \psi^\mathcal{I}_\mathcal{T}(X^\star, \bar{x}) \text{ for every } s \text{ with } s(\bar{x}) \in X^\star\\
\iff &(\mathfrak{A}, X) \models_s \psi^\star(X, \bar{x}) \text{ for every } s \text{ with } s(\bar{x}) \in X^\star \quad \text{(Claim 13)}\\
\iff &(\mathfrak{A}, X) \models_s \psi^\star(X, \bar{x}) \text{ for every } s \text{ with } (\mathfrak{A}, X) \models_s e_T(\bar{u}) \wedge X\bar{v} \quad \text{(due to (2))}\\
\iff &(\mathfrak{A}, X) \models \forall \bar{u} \forall \bar{v} \forall \bar{w}((e_T(\bar{u}) \wedge X\bar{v}) \rightarrow \psi^\star(X, \bar{u}, \bar{v}, \bar{w}))) \equiv \mu.
\end{aligned}$$

As a result, we have that $(\mathfrak{A}, X) \models \mu(X) \iff (\mathfrak{A}, X^\star) \models \varphi^\mathcal{I}_\mathcal{T}$. Putting everything together yields:

$$(\mathfrak{A}, X) \models \mu \iff (\mathfrak{A}, X^\star) \models \varphi^\mathcal{I}_\mathcal{T} \overset{(1)}{\iff} (\mathcal{G}, X) \models \varphi_\mathcal{T} \overset{\text{(Claim 12)}}{\iff} X \in \mathcal{T}(\mathcal{G}) \overset{\text{(Theorem 6)}}{\iff} (\mathfrak{A}, X) \models \varphi$$

Thus, the constructed myopic formula $\mu(X)$ is indeed equivalent to $\varphi(X)$. ◀

This construction can be applied to non union closed formulae as well, in which case the statement becomes $(\mathfrak{A}, \bigcup_{i \in I} X_i) \models \mu \iff (\mathfrak{A}, X_i) \models \varphi$ for all $i \in I$. To see this replace Claim 12 by "For every $X \subseteq A^r$, $(\mathcal{G}, X) \models \varphi_\mathcal{T}(X) \iff X = \bigcup_i X_i$ for some $X_i \in \mathcal{T}(\mathcal{G})$".

5 Union Games

In the previous section we have characterised the union closed fragment of Σ^1_1 by means of a syntactic normal form. Now we aim at a game theoretic description, which leads to the following restriction of inclusion-exclusion games that reveals *how* union closed properties are assembled.

▶ **Definition 14.** *A* union game *is an inclusion-exclusion game* $\mathcal{G} = (V, V_0, V_1, E, I, T, E_{\mathsf{ex}})$ *obeying the following restrictions. For every* $t \in T$ *the subgraph reachable from* t *via the edges* $E \setminus E_{\mathsf{in}}$*, that are the edges of* E *that do* not *go back into* T*, is denoted by* $\mathcal{G}^\triangle_t$.[4] *These*

[4] Recall that $E_{\mathsf{in}} := E \cap (V \times T)$.

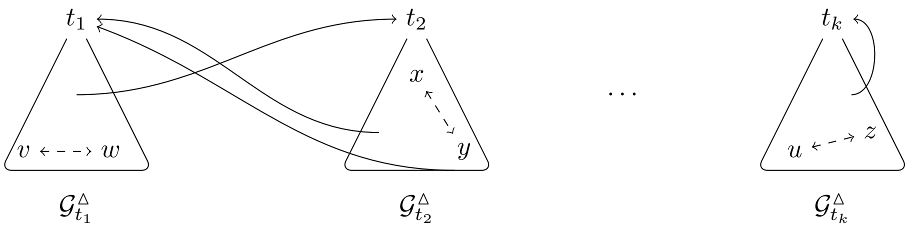

Figure 1 A drawing of a union game. The target positions $T = \{t_1, \dots, t_k\}$ are at the top of the components $\mathcal{G}^{\vartriangle}_t$, which are depicted by triangles. Recall that the inclusion edges, that are the edges going into target vertices, do not account for the reachability of the components $\mathcal{G}^{\vartriangle}_t$. The exclusion edges E_{ex} are drawn as dashed arrows and, as seen here, are allowed only inside a component.

components must be disjoint, that is $V(\mathcal{G}^{\vartriangle}_t) \cap V(\mathcal{G}^{\vartriangle}_{t'}) = \varnothing$ for all $t \neq t' \in T$. Furthermore, exclusion edges are only allowed between vertices of the same component, that is $E_{\mathsf{ex}} \subseteq \bigcup_{t \in T} V(\mathcal{G}^{\vartriangle}_t) \times V(\mathcal{G}^{\vartriangle}_t)$. The set of initial positions is empty, i.e. $I = \varnothing$.

See Figure 1 for a graphical representation of a union game. Since the exclusion edges are only inside a component we can in a way combine different strategies into one, which is the reason the target set of a union game is closed under unions.

▶ **Theorem 15.** *Let $\mathcal{G}$ be a union game and $(\mathcal{S}_i)_{i \in J}$ be a family of winning strategies for player 0. Then there is a winning strategy $\mathcal{S}$ for player 0 such that $\mathcal{T}(\mathcal{S}) = \bigcup_{i \in J} \mathcal{T}(\mathcal{S}_i)$. In other words, the set $\mathcal{T}(\mathcal{G})$ is closed under unions.*

Proof. Let $\mathcal{S}_i = (W_i, F_i)$ for $i \in J$. Let $U := \bigcup_{i \in J} \mathcal{T}(\mathcal{S}_i)$ and $f \colon U \to J$ be a function such that $t \in \mathcal{T}(\mathcal{S}_{f(t)})$ for all $t \in U$. Define $\mathcal{S} := \bigcup_{t \in U} (\mathcal{S}_{f(t)} {\restriction}_{V(\mathcal{G}^{\vartriangle}_t)} + (E(\mathcal{S}_{f(t)}) \cap (V(\mathcal{G}^{\vartriangle}_t) \times T)))$. In words, $\mathcal{S}$ is defined on every component $\mathcal{G}^{\vartriangle}_t$ with $t \in U$ as an arbitrary strategy $\mathcal{S}_t$ that is defined on $\mathcal{G}^{\vartriangle}_t$, including the inclusion edges leaving this component. By definition $\mathcal{T}(\mathcal{S}) = U$ and, furthermore, $\mathcal{S}$ is indeed a winning strategy since it behaves on every component $\mathcal{G}^{\vartriangle}_t$ like $\mathcal{S}_{f(t)}$ and there are no exclusion edges between different components. ◀

▶ **Definition 16.** *Let $\mu(X) = \forall \bar{x}(X\bar{x} \to \exists \bar{R}\varphi(X, \bar{R}, \bar{x}))$ be a myopic τ-formula where φ is in negation-normal form and $\mathfrak{A}$ be a τ-structure. The union game $\mathcal{G}(\mathfrak{A}, \mu) := (V, V_0, V_1, E, I = \varnothing, T = A^{\mathrm{ar}(X)}, E_{\mathsf{ex}})$ is defined similarly to Definition 5 with the difference being that for each $\bar{a} \in A^{\mathrm{ar}(\bar{x})}$ we have to play on a copy of the game, so positions are now of the form $(\vartheta, s, \bar{a})$ instead of (ϑ, s), where $\vartheta \in \mathrm{subf}(\varphi)$. The target vertices are the roots of these components, which is reflected by edges from $\bar{a}$ to $(\varphi, \bar{x} \mapsto \bar{a}, \bar{a})$. Because of this construction exclusion edges can only occur inside a component.*

Notice that there are still edges from $(X\bar{x}, s, \bar{a})$ to $s(\bar{x})$ – the inclusion edges. It is also worth mentioning that the empty set is always included in $\mathcal{T}(\mathcal{G}(\mathfrak{A}, \mu))$ for all myopic μ because $(\varnothing, \varnothing)$ is a (trivial) winning strategy for player 0. This mimics the behaviour that in case $X = \varnothing$, the formula $\forall \bar{x}(X\bar{x} \to \psi)$ is satisfied regardless of everything else. The analogue of Theorem 6 holds for union games and myopic formulae.

▶ **Proposition 17.** *Let $\mathfrak{A}$, μ and $\mathcal{G}(\mathfrak{A}, \mu)$ be as in Definition 16. Then $(\mathfrak{A}, X) \models \mu \iff X \in \mathcal{T}(\mathcal{G}(\mathfrak{A}, \mu))$.*

We want to end this section with the remark that for other fragments with certain closure properties natural restrictions of inclusion-exclusion games exist. Especially, forbidding exclusion edges at all leads to model-checking games for inclusion logic, while forbidding inclusion edges results in games suited for exclusion logic.

6 Myopic Fragment of Inclusion-Exclusion Logic

Similarly to the normal form of union closed Σ^1_1-formulae (see Section 4) we present syntactic restrictions of inclusion-exclusion logic $\mathrm{FO}(\subseteq, \mid)$ that correspond precisely to the union closed fragment $\mathcal{U}$[5]. Analogously to myopic Σ^1_1-formulae we will also present a normal form for all union closed $\mathrm{FO}(\subseteq, \mid)$-formulae.

▶ **Definition 18.** *A formula $\varphi(\bar{x}) \in \mathrm{FO}(\subseteq, \mid)$ is $\bar{x}$-myopic, if the following conditions are satisfied:*

(a) *The variables from $\bar{x}$ are never quantified in φ.*

(b) *Every exclusion atom occurring in φ is of the form $\bar{x}\bar{y} \mid \bar{x}\bar{z}$.*

(c) *Every inclusion atom occurring in φ is of the form $\bar{x}\bar{y} \subseteq \bar{x}\bar{z}$ or $\bar{y} \subseteq \bar{x}$, where the latter is* only *allowed if it is not in the scope of a disjunction.*

*Please note that $\varphi(\bar{x})$ must not have any additional free variables besides $\bar{x}$. We call atoms of the form $\bar{x}\bar{y} \subseteq \bar{x}\bar{z}$ or $\bar{x}\bar{y} \mid \bar{x}\bar{z}$ ($\bar{x}$-)*guarded *and $\bar{y} \subseteq \bar{z}$, respectively $\bar{y} \mid \bar{z}$, the corresponding* unguarded *versions. Analogously, we call a formula ψ the unguarded version of φ, if ψ emerges from φ by replacing every dependency atom by the respective unguarded version.*

The intuition behind this definition is that every $\bar{x}$-myopic formula can be evaluated componentwise on every team $X\restriction_{\bar{x}=\bar{a}} = \{s \in X : s(\bar{x}) = \bar{a}\}$ for all $\bar{a} \in X(\bar{x})$. For a formula φ let T_φ denote its syntax tree[6]. A (team-)labelling of T_φ is a function λ mapping every node v_ψ to a team $\lambda(v_\psi)$ whose domain includes $\mathrm{free}(\psi)$. In the following we write $\lambda(\psi)$ instead of $\lambda(v_\psi)$ if it is clear from the context which *occurrence* of the subformula ψ of φ is meant. We call λ a *witness* for $\mathfrak{A} \models_X \varphi$, if $\lambda(\varphi) = X$ and the semantical rules of Definition 1 are satisfied (e.g. $\lambda(\psi \vee \vartheta) = \lambda(\psi) \cup \lambda(\vartheta)$) and for every literal β of φ we have $\mathfrak{A} \models_{\lambda(\beta)} \beta$. By induction, if λ is a witness for $\mathfrak{A} \models_X \varphi$, then for every $\psi \in \mathrm{subf}(\varphi)$ we have $\mathfrak{A} \models_{\lambda(\psi)} \psi$ and, moreover, $\mathfrak{A} \models_X \varphi$ if and only if there is a witness λ for $\mathfrak{A} \models_X \varphi$.

▶ **Proposition 19.** *Let X be team over $\mathfrak{A}$ with $\mathrm{dom}(X) \supseteq \{\bar{x}, \bar{v}, \bar{w}\}$ and $\varphi(\bar{x})$ be $\bar{x}$-myopic.*

1. $\mathfrak{A} \models_X \bar{x}\bar{v} \subseteq \bar{x}\bar{w} \iff \mathfrak{A} \models_{X\restriction_{\bar{x}=\bar{a}}} \bar{v} \subseteq \bar{w}$ *for all $\bar{a} \in X(\bar{x})$*
2. $\mathfrak{A} \models_X \bar{x}\bar{v} \mid \bar{x}\bar{w} \iff \mathfrak{A} \models_{X\restriction_{\bar{x}=\bar{a}}} \bar{v} \mid \bar{w}$ *for all $\bar{a} \in X(\bar{x})$*
3. *For every subformula $\bar{v} \subseteq \bar{x}$ of φ and witness λ for $\mathfrak{A} \models_X \varphi$ we have $(\lambda(\bar{v} \subseteq \bar{x}))(\bar{x}) = X(\bar{x})$.*

Like union games an $\bar{x}$-myopic formula is evaluated componentwise, which leads to the union closure of this fragment.

▶ **Theorem 20.** *Let $\varphi(\bar{x}) \in \mathrm{FO}(\subseteq, \mid)$ be $\bar{x}$-myopic and $\mathfrak{A} \models_{X_i} \varphi$ for all $i \in I$. Then $\mathfrak{A} \models_X \varphi$ for $X = \bigcup_{i \in I} X_i$.*

It remains to prove that indeed every union closed formula φ of $\mathrm{FO}(\subseteq, \mid)$ is equivalent to some $\bar{x}$-myopic formula. As we have already seen in Theorem 8, every union closed formula of existential second-order logic is equivalent to some myopic Σ^1_1-formula. Moreover, it is well known that every $\mathrm{FO}(\subseteq, \mid)$-formula can be translated into an equivalent Σ^1_1-formula [1]. Such a formula can be expressed as an $\bar{x}$-myopic one of the form $\exists\bar{s}(\bar{s} \subseteq \bar{x} \wedge \psi)$ where ψ uses only $\bar{x}$-guarded atoms.

---

[5] We have defined $\mathcal{U}$ to be the set of all union closed Σ^1_1-formulae, by slight abuse of notation we use the same symbol here to denote the set of all $\mathrm{FO}(\subseteq, \mid)$-formulae that are closed under unions.

[6] Since we consider a tree instead of a DAG, identical subformulae may occur at different nodes.

▶ **Lemma 21.** *Let $\varphi(\bar{x}) \in \mathrm{FO}(\subseteq, \mid)$ be an $\bar{x}$-myopic formula of the form $\exists \bar{s}(\bar{s} \subseteq \bar{x} \wedge \psi)$, where in ψ no inclusion atoms of the form $\bar{y} \subseteq \bar{x}$ occur. Then $\mathfrak{A} \models_X \varphi$ if and only if there exists $F \colon X \to \mathcal{P}^+(A^{|\bar{x}|})$ such that $F(s) \subseteq X(\bar{x})$ for every $s \in X$ and $\mathfrak{A} \models_{X[\bar{s} \mapsto F]\restriction_{\bar{x}=\bar{a}}} \psi'$ for all $\bar{a} \in X(\bar{x})$, where ψ' is the unguarded version of ψ.*

Proof. By induction on ψ and applying Proposition 19. ◀

We present two different proofs for the next theorem, which bring a myopic Σ^1_1-formula into this normal form. The following proof is based on methods of Galliani, Kontinen and Väänänen [1, 8] while the other one resembles the proof of Theorem 11 and can be found in the full version.

▶ **Theorem 22.** *Let $\varphi(X)$ be a myopic Σ^1_1-formula. There is an equivalent $\bar{x}$-myopic formula of $\mathrm{FO}(\subseteq, \mid)$ where $|\bar{x}| = \mathrm{ar}(X)$.*

Proof. First of all let us introduce a normal form of myopic Σ^1_1-formulae. Since in myopic formulae the variable X may occur only positively in the subformula φ', we can transform every $\forall \bar{x}(X\bar{x} \to \exists \bar{R}\varphi'(\bar{R}, X, \bar{x}))$ into the equivalent formula $\forall \bar{x}(X\bar{x} \to \exists S(S \subseteq X \wedge \exists \bar{R}\varphi'(\bar{R}, S, \bar{x})))$, where $S \subseteq X$ is a shorthand for $\forall \bar{y}(S\bar{y} \to X\bar{y})$. We now apply the Skolem-normal form of Σ^1_1-formulae to $\exists \bar{R}\varphi'(\bar{R}, S, \bar{x})$, which yields the formula $\sigma(S, \bar{x}) \coloneqq \exists \bar{f}\forall \bar{y}((f_1(\bar{w}) = f_2(\bar{w}) \leftrightarrow S\bar{w}) \wedge \psi(\bar{f}, \bar{x}, \bar{y}))$, where ψ is a quantifier-free first-order formula and $\bar{w}$ is a subtuple of $\bar{y}$ and, moreover, every f_i occurs in σ only with a unique tuple $\bar{w}_i$ (consisting of pairwise different variables) as argument, that is $f_i(\bar{w}_i)$ (see [8] where an analogous construction is made). The original formula can thus be transformed into $\forall \bar{x}(X\bar{x} \to \exists S(S \subseteq X \wedge \sigma(S, \bar{x})))$. Similarly to [1] we embed $\sigma(S, \bar{x})$ into inclusion-exclusion logic as $\vartheta(\bar{s}, \bar{x}) \coloneqq \forall \bar{y}\exists \bar{z}\big(\bigwedge_i {=}(\bar{x}\bar{w}_i, z_i) \wedge ((\bar{x}\bar{w} \subseteq \bar{x}\bar{s} \wedge z_1 = z_2) \vee (\bar{x}\bar{w} \mid \bar{x}\bar{s} \wedge z_1 \neq z_2)) \wedge \psi'(\bar{x}, \bar{y}, \bar{z})\big)$. Here ψ' is obtained from ψ by simply replacing every occurrence of $f_i(\bar{w}_i) = f_j(\bar{w}_j)$ by $z_i = z_j$. The only difference in our case is that every dependency atom is $\bar{x}$-guarded due to the fact that the subformula at hand is inside the scope of the universally quantified variables $\bar{x}$ in $\forall \bar{x}(X\bar{x} \to \dots)$. Notice that dependence atoms of the form ${=}(\bar{x}\bar{w}_i, z_i)$ can also be regarded as $\bar{x}$-myopic. Formally, we can embed such an atom into exclusion logic via the formula $\forall v(\bar{x}\bar{w}_i v \mid \bar{x}\bar{w}_i z_i \vee z_i = v)$, which has the intended shape [1]. The whole formula $\varphi(X)$ thus translates into $\mu(\bar{x}) \coloneqq \exists \bar{s}(\bar{s} \subseteq \bar{x} \wedge \vartheta(\bar{s}, \bar{x}))$. Let $\vartheta'(\bar{s}, \bar{x})$ be the unguarded version of $\vartheta(\bar{s}, \bar{x})$. Analogously to the argumentation of Galliani [1] by additionally making use of Proposition 19, we see that $(\mathfrak{A}, Y\restriction_{\bar{x}=\bar{a}}(\bar{s})) \models_{\bar{x} \mapsto \bar{a}} \sigma(S, \bar{x})$ if and only if $\mathfrak{A} \models_{Y\restriction_{\bar{x}=\bar{a}}} \vartheta'(\bar{x})$ for $\bar{a} \in Y(\bar{x})$, where Y is a team with domain $\{\bar{s}, \bar{x}\}$ (here the variable S takes the role of the team). Using Lemma 21 we have $\mathfrak{A} \models_X \mu(\bar{x})$ if and only if there is a function $F \colon X \to \mathcal{P}^+(A^{\mathrm{ar}(\bar{s})})$ such that $F(s) \subseteq X(\bar{x})$ for every $s \in X$ and $\mathfrak{A} \models_{X[\bar{s} \mapsto F]\restriction_{\bar{x}=\bar{a}}} \vartheta'(\bar{s}, \bar{x})$ for all $\bar{a} \in X(\bar{x})$, which again holds if and only if there exists such an F and $(\mathfrak{A}, F(s)) \models_t \sigma(S, \bar{x})$ for all $t \in X$, but this just means $(\mathfrak{A}, X(\bar{x})) \models \forall \bar{x}(X\bar{x} \to \exists S(S \subseteq X \wedge \sigma(S, \bar{x})))$. ◀

▶ **Corollary 23** (Normal form of myopic-$\mathrm{FO}(\subseteq, \mid)$). *Let $\varphi(\bar{x})$ be a union closed formula of $\mathrm{FO}(\subseteq, \mid)$. There is a logically equivalent $\bar{x}$-myopic formula $\psi(\bar{x}) = \exists \bar{s}(\bar{s} \subseteq \bar{x} \wedge \vartheta)$ where in ϑ only $\bar{x}$-guarded dependency atoms occur.*

6.1 Optimality of the Myopic Fragment of Inclusion-Exclusion Logic

One might ask whether the restrictions of Definition 18 are actually imperative to capture the union closed fragment. In this section, we will show that neither condition can be dropped and that every single atom of Definition 18 is required to express all union closed properties.

We start by showing that neither condition can be dropped. First of all, it is pretty clear that exclusion atoms have to be $\bar{x}$-guarded, because $x_1 \mid x_2$ is not guarded and obviously

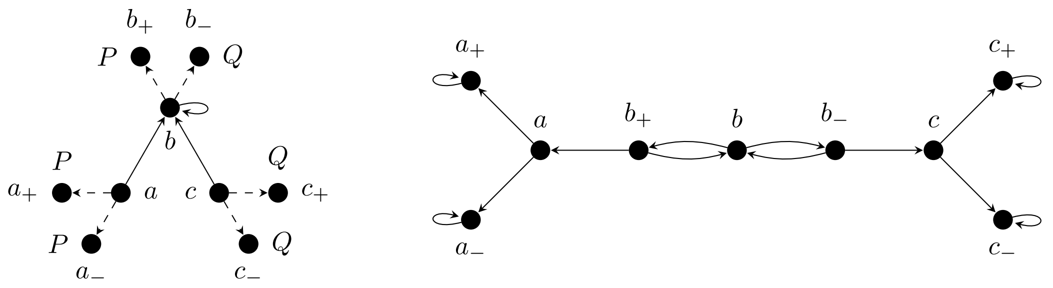

Figure 2 The structures $\mathfrak{A}$ and $\mathfrak{B}$. The structure $\mathfrak{A} = (V, E^{\mathfrak{A}}, F^{\mathfrak{A}}, P^{\mathfrak{A}}, Q^{\mathfrak{A}})$ on the left side uses two different kinds of edges: the dashed edges belong to F, while the other are E-edges. Furthermore, $\mathfrak{A}$ exhibits two predicates P, Q. The structure $\mathfrak{B} = (V, E^{\mathfrak{B}})$ depicted on the right is just a directed graph. Please notice that both structures are using the same universe V.

not closed under unions. Furthermore, it is clear that the variables among $\bar{x}$ must not be quantified. This points out the necessity of conditions (a) and (b) of Definition 18. In the next example we demonstrate that neither restriction of condition (c) can be dropped.

▶ **Example 24.** Consider the structures $\mathfrak{A}$ and $\mathfrak{B}$ drawn in Figure 2 and the following formulae:

$$\begin{aligned}
\varphi(x) &\coloneqq \exists y \exists z (Fxy \wedge Fxz \wedge xy \mid xz \wedge [(Py \wedge \vartheta(x)) \vee (Qy \wedge \vartheta(x))]) \\
&\qquad \text{where } \vartheta(x) \coloneqq \exists v (Exv \wedge v \subseteq x) \\
\psi(x) &\coloneqq \exists y \exists z (Exy \wedge Exz \wedge xy \mid xz \wedge \exists w (Eyw \wedge x \subseteq w))
\end{aligned}$$

Neither $\varphi(x)$ nor $\psi(x)$ is x-myopic, because the inclusion atom $v \subseteq x$ from ϑ occurs inside the scope of a disjunction (and it is not x-guarded), while the atom $x \subseteq w$ is neither x-guarded nor of the form that is allowed outside the scope of disjunctions, because x appears on the wrong side of the inclusion atom.

For every $v \in V$ let $s_v \colon \{x\} \to V$ be the assignment with $s_v(x) \coloneqq v$. We define the teams $X_1 \coloneqq \{s_a, s_b\}$, $X_2 \coloneqq \{s_b, s_c\}$ and $X \coloneqq X_1 \cup X_2 = \{s_a, s_b, s_c\}$. It is not difficult to verify that $\mathfrak{A} \vDash_{X_i} \varphi(x)$ and $\mathfrak{B} \vDash_{X_i} \psi(x)$ for $i = 1, 2$ but $\mathfrak{A} \nvDash_X \varphi(x)$ and $\mathfrak{B} \nvDash_X \psi(x)$. In particular, neither $\varphi(x)$ nor $\psi(x)$ is closed under unions. This shows that the restrictions of Definition 18 are indeed necessary.

Thus the atoms allowed in Definition 18 are sufficient to capture the union closed fragment of FO($\subseteq$, $\mid$). On the contrary, one may ask whether the set of atoms given in Definition 18 is necessary. Let us argue for all rules of Definition 18.

Assume that all exclusion atoms are forbidden. Then every formula is already in inclusion logic in which one cannot define every union closed property as was shown by Galliani and Hella [2, p. 16].

If inclusion atoms were only allowed in the form $\bar{x}\bar{y} \subseteq \bar{x}\bar{z}$, that means the atoms $\bar{y} \subseteq \bar{x}$ are forbidden, the formulae become flat, as can be seen by considering Proposition 19, but not all union closed properties are flat.

The case where inclusion atoms of form $\bar{x}\bar{y} \subseteq \bar{x}\bar{z}$ are forbidden is a bit more delicate. To prove that such a formula cannot express every union closed property consider the formula $\mu(x) = \exists z (z \subseteq x \wedge \forall y (Exy \to xy \subseteq xz))$, where $\tau = \{E\}$ for a binary predicate symbol E. This formula axiomatises the set of all teams X over a graph $G = (V, E)$ such that whenever $v \in X(x)$ and $(v, w) \in E$, then already $w \in X(x)$. The formula obviously describes a union closed property. Consider the graph G: $b \longleftarrow a \longrightarrow c$. Here, $G \vDash_X \mu(x)$ for precisely those teams X that satisfy "$a \in X(x)$ implies $b, c \in X(x)$". For every $v \in V(G)$ let s_v be the assignment $x \mapsto v$ and let $X_v \coloneqq \{s_v\}$. Furthermore, we define $X_{abc} \coloneqq \{s_a, s_b, s_c\}$.

Let $\psi(x)$ be an x-myopic formula in which the construct $x\bar{y} \subseteq x\bar{z}$ does not appear. So the only inclusion atoms occurring in $\psi(x)$ are of the form $z \subseteq x$, which are not allowed in the scope of disjunctions. Notice that z cannot be universally quantified, as the team $X_b = \{s_b\}$ satisfies the described property, but not $\forall z(z \subseteq x)$. Thus we may assume without loss of generality that $\psi(x)$ has the form $\exists z(z \subseteq x \wedge \psi'(x, z))$, where in $\psi'(x, z)$ no atom of the kind $z' \subseteq x$ occurs. We want to remark that the following argumentation can be adapted to the slightly more general case that multiple atoms of form $z \subseteq x$ occur, but for sake of simplicity we only deal with one such atom. Let $\eta(x, z)$ be the unguarded version of $\psi'(x, z)$. By Lemma 21, there is a function $F\colon X_{abc} \to \mathcal{P}^+(V(G))$ such that $F(s) \subseteq X_{abc}(x) = V(G)$ for $s \in X_{abc}$ and $G \vDash_{X_{abc}[z \mapsto F]\upharpoonright_{x=v}} \eta$ for every $v \in X_{abc}(x)$. Please notice that $X_{abc}[z \mapsto F]\upharpoonright_{x=v} = X_v[z \mapsto F]$. Moreover, because in $\eta(x, z)$ no inclusion atom occurs it is downwards closed. Assume $a \in F(s_a)$. By downwards closure of $\eta(x, z)$ we obtain $G \vDash_{X_a[z \mapsto a]} \eta$, which, by Lemma 21, implies that $G \vDash_{X_a} \psi$ contradicting our assumption that ψ describes the desired property. Otherwise, because of symmetry, b is in $F(s_a)$, and hence $G \vDash_{X_a[z \mapsto b]} \eta$. Additionally, since $G \vDash_{X_b} \psi$ we know, by Lemma 21, that $G \vDash_{X_b[z \mapsto b]} \eta$. Together this implies $G \vDash_{X_{ab}[z \mapsto b]\upharpoonright_{x=v}} \eta$ for $v = a, b$ and, due to Lemma 21, we get $G \vDash_{X_{ab}} \psi$ which is again in conflict with our assumption about ψ describing the desired property.

7 An Atom capturing the Union Closed Fragment

The present work was motivated by a question of Galliani and Hella in 2013 [2]. Galliani and Hella asked whether there is a union closed atomic dependency notion α that is definable in existential second-order logic such that $\mathrm{FO}(\alpha)$ corresponds precisely to all union closed properties of $\mathrm{FO}(\subseteq, \mid)$. In [2] they have already shown that inclusion logic does not suffice, as there are union closed properties not definable in it. Moreover, they have established a theorem stating that every union closed atomic property that is definable in first-order logic (where the formula has access to the team via a predicate) is expressible in inclusion logic. Thus, whatever atom characterises all union closed properties of $\mathrm{FO}(\subseteq, \mid)$ must axiomatise an inherently second-order property.

Intuitively speaking, as we have seen in Section 5, solving union games is a complete problem for the class $\mathcal{U}$. Therefore, a canonical solution to this question is to propose an atomic formula that defines the winning regions in a union game. Towards this we must describe how a game can be encoded into a team. This is not as straightforward as one might think, because there is a technical pitfall we need to avoid. The union of two teams describing union games, each won by player 0, might encode a game won by player 1, but by union closure it must satisfy the atomic formula.

We encode union games in teams by using variable tuples for the respective components, where we also encode the complementary relations in order to ensure that the union of two different games cannot form a different game. For $k \in \mathbb{N}$ let $\mathcal{V}_k$ be the set of distinct k-tuples of variables $\{\bar{u}, \bar{v}_0, \bar{v}_1, \bar{v}, \bar{w}, \bar{t}, \bar{v}_{\mathsf{ex}}, \bar{w}_{\mathsf{ex}}, \bar{\varepsilon}_1, \bar{\varepsilon}_2, \bar{u}^{\mathsf{C}}, \bar{v}^{\mathsf{C}}, \bar{w}^{\mathsf{C}}, \bar{t}^{\mathsf{C}}, \bar{v}_{\mathsf{ex}}^{\mathsf{C}}, \bar{w}_{\mathsf{ex}}^{\mathsf{C}}, \bar{\varepsilon}_1^{\mathsf{C}}, \bar{\varepsilon}_2^{\mathsf{C}}\}$.

▶ **Definition 25.** *Let X be a team with $\mathcal{V}_k \subseteq \mathrm{dom}(X)$ and codomain A. We define $\sim \coloneqq X(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$ and $\mathfrak{A}^X \coloneqq (V, V_0, V_1, E, I, T, E_{\mathsf{ex}})$ with the following components.*

- $V \coloneqq X(\bar{u})$
- $V_0 \coloneqq X(\bar{v}_0)$
- $V_1 \coloneqq X(\bar{v}_1)$
- $E \coloneqq X(\bar{v}, \bar{w})$
- $I \coloneqq \varnothing$
- $T \coloneqq X(\bar{t})$
- $E_{\mathsf{ex}} \coloneqq X(\bar{v}_{\mathsf{ex}}, \bar{w}_{\mathsf{ex}})$

If the following consistency requirements are satisfied, then we define $\mathcal{G}^A_X \coloneqq \mathfrak{A}^X_{/\sim}$*.*

1. $X(\bar{u}^{\mathsf{C}}) = A^k \setminus V$
2. $X(\bar{v}^{\mathsf{C}}, \bar{w}^{\mathsf{C}}) = (A^k \times A^k) \setminus E$
3. $X(\bar{t}^{\mathsf{C}}) = A^k \setminus T$
4. $X(\bar{v}^{\mathsf{C}}_{\mathsf{ex}}, \bar{w}^{\mathsf{C}}_{\mathsf{ex}}) = (A^k \times A^k) \setminus E_{\mathsf{ex}}$
5. $X(\bar{\varepsilon}^{\mathsf{C}}_1, \bar{\varepsilon}^{\mathsf{C}}_2) = (A^k \times A^k) \setminus \sim$
6. $V_0 = V \setminus V_1$
7. $\mathfrak{A}^X$ *is a structure*[7]*.*
8. $\sim$ *is a congruence on* $\mathfrak{A}^X$*.*
9. $\mathfrak{A}^X_{/\sim}$ *is a union game.*

Otherwise, if any of these requirements is not fulfilled, we let $\mathcal{G}^A_X$ *be undefined.*

We call X complete (w.r.t. A), if $X(\bar{y}) \cup X(\bar{y}^{\mathsf{C}})$ is A^k or $A^k \times A^k$ for every $\bar{y} \in \{(\bar{u}), (\bar{v}, \bar{w}), (\bar{t}), (\bar{v}_{\mathsf{ex}}, \bar{w}_{\mathsf{ex}}), (\bar{\varepsilon}_1, \bar{\varepsilon}_2)\}$ and $V = V_0 \cup V_1$, and incomplete otherwise. It is easy to observe that $\mathcal{G}^A_X$ is undefined for every incomplete team X. Furthermore complete subteams of teams describing a game actually describe the same game and the same congruence relation.

▶ **Lemma 26.** *Let X, Y be teams with codomain A and $\mathcal{V}_k \subseteq \mathrm{dom}(X) = \mathrm{dom}(Y)$. If X is complete, $X \subseteq Y$ and $\mathcal{G}^A_Y$ is defined, then $\mathcal{G}^A_X = \mathcal{G}^A_Y$ and $\sim_X \coloneqq X(\bar{\varepsilon}_1, \bar{\varepsilon}_2) = Y(\bar{\varepsilon}_1, \bar{\varepsilon}_2) =: \sim_Y$.*

Now let us show that union games are definable in plain first-order logic with team semantics in the sense of Definition 25.

▶ **Lemma 27.** *Let $\varphi(X) = \forall \bar{x}(X\bar{x} \to \exists \bar{R} \varphi'(X, \bar{R}, \bar{x}))$ be a myopic Σ^1_1-formula and $\psi(\mathcal{V}_k, \bar{x})$ be a formula with team semantics (where k is large enough such that the game $\mathcal{G}(\mathfrak{A}, \varphi)$ can be encoded). There is a formula $\vartheta^\psi_\varphi(\bar{x})$ such that $\mathfrak{A} \vDash_X \vartheta^\psi_\varphi \iff \mathfrak{A} \vDash_Y \psi$ for some team Y extending X with $\mathcal{G}^A_Y \cong \mathcal{G}(\mathfrak{A}, \varphi)$ and $X(\bar{x}) = Y(\bar{x})$, for every τ-structure $\mathfrak{A}$.*

Proof. Similar to Lemma 10, it is easy to construct a (quantifier-free) first-order interpretation $\mathcal{I} \coloneqq (\delta, \varepsilon, \psi_V, \psi_{V_0}, \psi_{V_1}, \psi_E, \psi_I, \psi_T, \psi_{E_{\mathsf{ex}}})$ with $\mathcal{I}(\mathfrak{A}) \cong \mathcal{G}(\mathfrak{A}, \varphi)$. Now let $\vartheta^\psi_\varphi(\bar{x}) \coloneqq \forall \mathcal{V}_k(\gamma(\mathcal{V}_k) \to \psi(\mathcal{V}_k, \bar{x}))$ where the formula

$$\begin{aligned} \gamma(\mathcal{V}_k) \coloneqq\ & \delta(\bar{u}) \wedge \psi_{V_0}(\bar{v}_0) \wedge \psi_{V_1}(\bar{v}_1) \wedge \psi_E(\bar{v}, \bar{w}) \wedge \psi_T(\bar{t}) \wedge \psi_{E_{\mathsf{ex}}}(\bar{v}_{\mathsf{ex}}, \bar{w}_{\mathsf{ex}}) \wedge \varepsilon(\bar{\varepsilon}_1, \bar{\varepsilon}_2) \wedge \\ & \neg\delta(\bar{u}^{\mathsf{C}}) \wedge \neg\psi_E(\bar{v}^{\mathsf{C}}, \bar{w}^{\mathsf{C}}) \wedge \neg\psi_T(\bar{t}^{\mathsf{C}}) \wedge \neg\psi_{E_{\mathsf{ex}}}(\bar{v}^{\mathsf{C}}_{\mathsf{ex}}, \bar{w}^{\mathsf{C}}_{\mathsf{ex}}) \wedge \neg\varepsilon(\bar{\varepsilon}^{\mathsf{C}}_1, \bar{\varepsilon}^{\mathsf{C}}_2) \end{aligned}$$

enforces that the game $\mathcal{G}(\mathfrak{A}, \varphi)$ will be "loaded" into the team. As long as none of these conjuncts are unsatisfiable this construction is correct. This is safe to assume because one can easily transform a union game into an equivalent one w.r.t. the target set such that none of its components are empty. ◀

This knowledge enables us to finally define the atomic formula we sought after. For this we need to show that the atom is union closed and its first-order closure can express all of $\mathcal{U}$.

▶ **Definition 28.** *The atomic team formula $\cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$ for the respective tuples of variables has the following semantics. For non-empty teams X with $\mathcal{V}_k, \bar{x} \subseteq \mathrm{dom}(X)$ we define*

$$\mathfrak{A} \vDash_X \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x}) :\iff X \textit{ is complete and if } \mathcal{G}^A_X \textit{ is defined, then } X(\bar{x})_{/X(\bar{\varepsilon}_1, \bar{\varepsilon}_2)} \in \mathcal{T}(\mathcal{G}^A_X)$$

and we set $\mathfrak{A} \vDash_\varnothing \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$ to be always true (to ensure the empty team property).

[7] This condition ensures that $V_0 \subseteq V$ and so forth.

Note that this atom can be defined in existential second-order logic.

▶ **Proposition 29.** *The atomic formula* $\cup\!-\!\mathsf{game}$ *is union closed.*

Proof. Assume that $\mathfrak{A} \models_{X_i} \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$ for $i \in I$. We prove that $\mathfrak{A} \models_X \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$ holds for the union $X := \bigcup_{i\in I} X_i$. If $X = \varnothing$, there is nothing to prove. Otherwise at least one X_j is non-empty and, since $\mathfrak{A} \models_{X_j} \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$, X_j must be complete implying that X is also complete (because $X \supseteq X_j$). For the remainder of this proof, we assume w.l.o.g. that all involved teams X_i (and X) are non-empty. If $\mathcal{G}^A_X$ is undefined, then $\mathfrak{A} \models_X \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$ follows from the definition of $\cup\!-\!\mathsf{game}$. Otherwise, if $\mathcal{G}^A_X$ is defined, then we can use Lemma 26 to obtain that $\mathcal{G}^A_X = \mathcal{G}^A_{X_i}$ and $\sim := X(\bar{\varepsilon}_1, \bar{\varepsilon}_2) = X_i(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$ for every $i \in I$. Since $\mathfrak{A} \models_{X_i} \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$, we can conclude that $X_i(\bar{x})_{/\sim} \in \mathcal{T}(\mathcal{G}^A_{X_i}) = \mathcal{T}(\mathcal{G}^A_X)$ for each $i \in I$. By Theorem 15, $X(\bar{x})_{/\sim} = \bigcup_{i\in I} X_i(\bar{x})_{/\sim} \in \mathcal{T}(\mathcal{G}^A_X)$ and, hence, $\mathfrak{A} \models_X \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{x})$. ◀

▶ **Theorem 30.** *Let* $\varphi \in \mathrm{FO}(\subseteq, \mid)$ *be a union closed formula. There is a logically equivalent formula* $\zeta \in \mathrm{FO}(\cup\!-\!\mathsf{game})$*. In other words,* $\mathrm{FO}(\cup\!-\!\mathsf{game})$ *corresponds precisely to the union closed fragment of* $\mathrm{FO}(\subseteq, \mid)$*.*

Proof. Let $\mathfrak{A}$ be an arbitrary structure. Due to [1, Theorem 6.1] there exists a formula $\varphi'(X) \in \Sigma^1_1$ which is logically equivalent to $\varphi(\bar{x})$ in the sense that $\mathfrak{A} \models_X \varphi(\bar{x}) \iff (\mathfrak{A}, X(\bar{x})) \models \varphi'(X)$ for every team X with $\bar{x} \subseteq \mathrm{dom}(X)$. By Theorem 8, there is a myopic formula $\mu \equiv \varphi'$. So, we have $(\mathfrak{A}, X(\bar{x})) \models \mu(X) \iff \mathfrak{A} \models_X \varphi(\bar{x})$.

The game $\mathcal{G}(\mathfrak{A}, \mu)$ from Definition 16 is a union game and Lemma 27 allows us to load this game into a team. Please notice, that Lemma 27 is using a similar first-order interpretation $\mathcal{I}$ as Lemma 10, which encodes a target vertex $\bar{a} \in T(\mathcal{G}(\mathfrak{A}, \mu))$ by tuples of the form $(\bar{u}, \bar{a}, \bar{w})$ of length $k = n + m$ where the n-tuple $\bar{u}$ has the equality type e_T while $\bar{w}$ is an arbitrary tuple of length $m - |\bar{a}|$. Let $\psi(\mathcal{V}_k, \bar{x}) := \forall\bar{u}\forall\bar{w}(e_T(\bar{u}) \to \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{u}\bar{x}\bar{w}))$ and $\zeta(\bar{x}) := \vartheta^\psi_\mu$ be as in Lemma 27, that is $\forall\mathcal{V}_k(\gamma(\mathcal{V}_k) \to \psi(\mathcal{V}_k, \bar{x}))$. So $\mathfrak{A} \models_X \zeta(\bar{x}) \iff \mathfrak{A} \models_Y \psi(\mathcal{V}_k, \bar{x})$ where $Y = X[\mathcal{V}_k \mapsto A]\restriction_\gamma$. As in Lemma 27, we have $\mathcal{G}^A_Y \cong \mathcal{I}(\mathfrak{A}) \cong \mathcal{G}(\mathfrak{A}, \mu)$ and $X(\bar{x}) = Y(\bar{x})$. Furthermore, we have defined $\mathcal{G}^A_Y = \mathfrak{A}^Y_{/\sim}$ where $\sim := Y(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$.

Because of the construction of ψ, we have $\mathfrak{A} \models_Y \psi(\mathcal{V}_k, \bar{x}) \iff \mathfrak{A} \models_Z \cup\!-\!\mathsf{game}(\mathcal{V}_k, \bar{u}\bar{x}\bar{w})$ where $Z := Y[\bar{u} \mapsto e^{\mathfrak{A}}_T, \bar{w} \mapsto A^{m-|\bar{x}|}]$. Since $\mathcal{G}^A_Z = \mathcal{G}^A_Y \cong \mathcal{G}(\mathfrak{A}, \mu)$ is a well-defined union game, this is equivalent to $Z(\bar{u}\bar{x}\bar{w})_{/\sim} \in \mathcal{T}(\mathcal{G}^A_Y)$. Let $h\colon \delta^{\mathfrak{A}}_{\mathcal{I}} \to V(\mathcal{G}(\mathfrak{A}, \mu))$ be the coordinate map for $\mathcal{G}(\mathfrak{A}, \mu) \cong \mathcal{I}(\mathfrak{A})$. By construction, h induces an isomorphism between $\mathfrak{A}^Y_{/\sim}$ and $\mathcal{G}(\mathfrak{A}, \mu)$. In particular each element of any equivalence class $[(\bar{u}', \bar{a}, \bar{w}')]_\sim \in Z(\bar{u}\bar{x}\bar{w})_{/\sim}$ is mapped by h to $\bar{a}$. Therefore, $Z(\bar{u}\bar{x}\bar{w})_{/\sim} \in \mathcal{T}(\mathcal{G}^A_Y) \iff Z(\bar{x}) = X(\bar{x}) \in \mathcal{T}(\mathcal{G}(\mathfrak{A}, \mu))$. Thus we have $\mathfrak{A} \models_X \zeta(\bar{x}) \iff X(\bar{x}) \in \mathcal{T}(\mathcal{G}(\mathfrak{A}, \mu))$. Putting everything together, we have $\mathfrak{A} \models_X \zeta(\bar{x}) \iff X(\bar{x}) \in \mathcal{T}(\mathcal{G}(\mathfrak{A}, \mu)) \iff (\mathfrak{A}, X(\bar{x})) \models \mu \iff \mathfrak{A} \models_X \varphi(\bar{x})$ as desired. ◀

8 Concluding Remarks

Let us remark on the "naturalness" of the atom $\cup\!-\!\mathsf{game}$. Certainly inclusion, exclusion and the notions alike can be regarded as natural atomic dependency formulae, whereas the just introduced atom has to be classified differently. Nevertheless, it is a canonical candidate since it solves a complete problem of the desired class. Of course, a more natural – and more usable – atom might be found, but it will not be as simplistic as e.g. inclusion for Galliani and Hella have shown that every first-order definable union closed property is already expressible in inclusion logic. Hence, whatever atom one proposes, it must make use of some inherently second-order concepts. For concretely expressing properties, the introduced myopic fragments of Σ^1_1 and $\mathrm{FO}(\subseteq, \mid)$ are more practical.

The various syntactical characterisations of the union closed fragments presented in this work now enables their further investigation. This could result in a complexity theoretical analysis or a more detailed classification of Σ^1_1.

References

1. Pietro Galliani. Inclusion and Exclusion Dependencies in Team Semantics—on Some Logics of Imperfect Information. *Annals of Pure and Applied Logic*, 163(1):68–84, 2012.
2. Pietro Galliani and Lauri Hella. Inclusion Logic and Fixed Point Logic. In Simona Ronchi Della Rocca, editor, *Computer Science Logic 2013 (CSL 2013)*, volume 23 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 281–295, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.CSL.2013.281`.
3. Erich Grädel. Games for Inclusion Logic and Fixed-Point Logic. In Samson Abramsky and Others, editors, *Dependence Logic: Theory and Applications*, pages 73–98. Birkhäuser, 2016.
4. Erich Grädel, Phokion G. Kolaitis, Leonid Libkin, Maarten Marx, Joel Spencer, Moshe Y. Vardi, Yde Venema, and Scott Weinstein. *Finite Model Theory and Its Applications*. Springer-Verlag, Berlin Heidelberg, 2007.
5. Erich Grädel and Jouko Väänänen. Dependence and Independence. *Studia Logica*, 101(2):399–410, 2013. `doi:10.1007/s11225-013-9479-2`.
6. Wilfrid Hodges. *A Shorter Model Theory*. Cambridge University Press, New York, NY, USA, 1997.
7. Wilfrid Hodges. Compositional Semantics for a Language of Imperfect Information. *Logic Journal of the IGPL*, 5(4):539–563, 1997. `doi:10.1093/jigpal/5.4.539`.
8. Juha Kontinen and Jouko Väänänen. On Definability in Dependence Logic. *Journal of Logic, Language and Information*, 18(3):317–332, 2009.
9. Jouko Väänänen. *Dependence Logic: A New Approach to Independence Friendly Logic (London Mathematical Society Student Texts)*. Cambridge University Press, New York, NY, USA, 2007.
10. Johan van Benthem. *Modal Correspondence Theory*. PhD dissertation, University of Amsterdam, 1976.