Partially ordered connectives and Σ_1^1 on finite models

Merlijn Sevenster[1,⋆] and Tero Tulenheimo[2,⋆⋆]

[1] Institute for Logic, Language and Computation, University of Amsterdam,
Plantage Muidergracht 24, 1018 TV, Amsterdam, The Netherlands
sevenstr@science.uva.nl

[2] Department of Philosophy, University of Helsinki,
P.O. Box 9 (Siltavuorenpenger 20 A), 00014 University of Helsinki, Finland
tero.tulenheimo@helsinki.fi

Abstract. In this paper we take up the study of Henkin quantifiers with boolean variables [4] also known as partially ordered connectives [19]. We consider first-order formulae prefixed by partially ordered connectives, denoted $\mathbf{D}$, on finite structures. We characterize $\mathbf{D}$ as a fragment of second-order existential logic $\Sigma_1^1\heartsuit$ whose formulae do not allow for existential variables being argument of predicate variables. We show that $\Sigma_1^1\heartsuit$ harbors a strict hierarchy induced by the arity of predicate variables and that it is not closed under complementation, by means of a game-theoretical argument. Admitting for at most one existential variable to appear as the argument of a predicate variable already yields a logic coinciding with full Σ_1^1, thus we show.
Keywords. Henkin quantifiers, partially ordered connectives, NP vs. coNP, finite model theory

1 Introduction

Fagin's Theorem [9], stating that NP $= \Sigma_1^1$, reveals the intimate connection between finite model theory and complexity theory. As a methodological consequence it appears that questions and results regarding a complexity class may bear relevance for logic and vice versa. For instance, the complexity theorist's headache caused by the NP = coNP-problem can now be shared by the logician working on the $\Sigma_1^1 = \Pi_1^1$-problem.[3] Indeed, logicians took up the challenge and nowadays separating logics related to complexity classes is one of their main occupations. By and large they go about by mapping out *fragments* of various relevant logics. A point in case is Fagin's [10] study of the *monadic* fragments of Σ_1^1 and Π_1^1, showing that they do not coincide.

The results in [10] did arouse a lot of interest in monadic languages [2, 3, 20], but somewhat disappointingly, we are still waiting for methods to separate

⋆ Corresponding author

⋆⋆ Supported by NWO visitor's grant B 62-608

[3] Solving the NP = coNP-problem is worth a headache indeed: if NP ≠ coNP then P ≠ NP.

binary, existential, second-order logic from 3-ary, existential, second-order logic, see [5], or even from binary, universal, second-order logic.

In the present paper we concern ourselves with the finite model theory of languages with *Henkin quantifiers with restricted quantifiers* also known as *partially ordered connectives*. Henkin quantifiers $\mathsf{H}^n_k \boldsymbol{xy}$ are objects of the form

$$\begin{pmatrix} \forall x_{11} \ldots \forall x_{1k} \ \exists y_1 \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ \forall x_{n1} \ldots \forall x_{nk} \ \exists y_n \end{pmatrix} \tag{1}$$

that prefix first-order formulae ϕ. Here and henceforth, a series of variables as in $x_{11}, \ldots, x_{nk}$ is abbreviated by $\boldsymbol{x}$. On suitable structures $\mathfrak{A}$, the formula $\mathsf{H}^n_k \boldsymbol{xy}\ \phi(\boldsymbol{x}, \boldsymbol{y})$ is defined to be true iff there are functions $f_1, \ldots, f_n$ such that

$$\mathfrak{A} \models \forall \boldsymbol{x}\ \phi(\boldsymbol{x}, f_1(\boldsymbol{x}_1), \ldots, f_n(\boldsymbol{x}_n)), \tag{2}$$

where $\boldsymbol{x}_i = x_{i1}, \ldots, x_{ik}$. It is a milestone result in the theory of Henkin quantification that the logic obtained by applying objects as in (1) to first-order formulae, denoted $\mathbf{H}$, coincides with Σ^1_1, cf. [8, 21]. Referring to Fagin's Theorem, Blass and Gurevich [4, Theorem 1] draw the conclusion that $\mathbf{H} = \Sigma^1_1 = \mathrm{NP}$. In the same publication the authors study what constraints can be imposed on the existentially quantified variables in a Henkin quantifier, such as $\boldsymbol{y}$ in (1), without the quantifier losing its power to express NP-complete problems. It turns out that Henkin quantifiers of the form

$$\begin{pmatrix} \forall x_{11} \ldots \forall x_{1k} \ \exists \alpha_1 \\ \forall x_{21} \ldots \forall x_{2k} \ \exists \alpha_2 \end{pmatrix} \tag{3}$$

cannot express NP-complete problems, unless $\mathrm{NL} = \mathrm{NP}$. The variables α_1 and α_2 appearing in (3) are *boolean variables* that range over a fixed domain $\{0, 1\}$, say. In this sense $\exists \alpha_i$ is a 'restricted' quantifier, hence the name 'Henkin quantifier with restricted quantifiers'.

The model theory for Henkin quantifiers with restricted variables was taken up in [19], be it under the name of 'partially ordered connectives' and written in the following format:

$$\begin{pmatrix} \forall x_{11} \ldots \forall x_{1k} \ \bigvee i_1 \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ \forall x_{n1} \ldots \forall x_{nk} \ \bigvee i_n \end{pmatrix}, \tag{4}$$

denoted $\mathsf{D}^n_k \boldsymbol{xi}$. The usage of the symbol $\bigvee$ reflects the fact that the variables i_j range over a fixed domain of $\{0, 1\}$. Sandu and Väänänen [19, Proposition 2] show that any first-order formula ϕ prefixed by the partially ordered connective $\mathsf{D}^2_1 \boldsymbol{xi}$ can be translated into $\mathsf{H}^2_1 \boldsymbol{xy}\ \phi'$, for some first-order ϕ'. Furthermore, they provide an *Ehrenfeucht-Fraïssé game* for partially ordered connectives and use it to give non-definability results. Note that there are first-order formulae ϕ, that can express NP-complete problems, when prefixed with the partially ordered connective $\mathsf{D}^3_1 \boldsymbol{x}\boldsymbol{i}$, in virtue of Blass and Gurevich's result; 3-colorability of graphs is a point in case.

Other publications on Henkin quantifiers and partially ordered connectives in relation with complexity theory include [13, 14, 16–18].

In this paper we characterize the logic $\mathbf{D}$ – the result of applying (4) to first-order formulae for arbitrary k, n – as a fragment of Σ^1_1. The relevant fragment of Σ^1_1 only allows for variables occurring as arguments of predicate variables that are universally quantified. As this constraint is rather natural it may be of interest to the descriptive complexity community to observe that (a) $\mathbf{D}$ can express a property that can be expressed in $k+1$-ary, existential, second-order logic that cannot be expressed in k-ary, existential, second-order logic and that (b) $\mathbf{D}$ is not closed under complementation, as it can express 2-Colorability but not its complement. On the fly we prove that the Henkin quantifier $\mathsf{H}^2_1 \boldsymbol{x}$ is not definable in $\mathbf{D}$ and that $\mathbf{D}$ is strictly contained in NP.

In Section 2, we introduce the necessary apparatus to get going. In Section 3, we characterize $\mathbf{D}$ as a fragment of Σ^1_1. Using this characterization we show result (a). In Section 4, we give an Ehrenfeucht-Fraïssé game for $\mathbf{D}$ and use it to show that $\mathbf{D}$ is not closed under complementation, cf. (b). In Section 5 we show that if we extend $\Sigma^1_1\heartsuit$ so as to allow for at most one existential variable to occur as argument of predicate variables, the resulting logic coincides with full Σ^1_1.

2 Preliminaries

A *vocabulary* τ is a finite set of relation symbols, rigidly including the equality symbol. Vocabularies do not contain constant of function symbols. Results can easily be extended to vocabularies with constant symbols, though. A *finite τ-structure* $\mathfrak{A} = \langle A, \langle R^{\mathfrak{A}} \rangle_{R\in\tau} \rangle$ consists of a finite set A, referred to as the *universe of* $\mathfrak{A}$, and interpretations of the relation symbols in τ on A. Here and henceforth, every structure is finite and for this reason we omit mentioning this. The equality symbol is interpreted as the identity relation. If τ only contains one binary relation symbol, other than the equality symbol, then any τ-structure is called a *directed graph (digraph).* If $\mathfrak{G} = \langle G, R^{\mathfrak{G}} \rangle$ is a digraph and $R^{\mathfrak{G}}$ is symmetric, then $\mathfrak{G}$ is a *graph.* A class relevant to this paper is n-Colorability holding of those finite graphs whose chromatic number is $\leq n$. Conversely, let $\overline{n\text{-Colorability}}$ denote the complement of n-Colorability with respect to the class of finite graphs.

Define an *implicit matrix τ-formula* γ as a function of type $\{0,1\}^k \to \mathbf{FO}(\tau)$, where k is an integer k and $\mathbf{FO}(\tau)$ is first-order logic over τ. Let $\mathbf{D}_k(\tau)$ be the logic with formulae of the form $\mathsf{D}^n_k \boldsymbol{x}\boldsymbol{i}\ \gamma(\boldsymbol{i})(\boldsymbol{x})$, for arbitrary n. The notion of *bound* and *free variable* is canonically extended from first-order logic so as to apply to the variables $\boldsymbol{i}$ as well. A *sentence* is a formula without free variables. We shall usually omit explicit indication of as many variables from the formulae as possible without losing on readability. In this manner we may write $\mathsf{D}^n_k\gamma$ instead of $\mathsf{D}^n_k \boldsymbol{x}\boldsymbol{i}\ \gamma(\boldsymbol{i})(\boldsymbol{x})$. Put $\mathbf{D}(\tau) = \bigcup_k \mathbf{D}_k(\tau)$.

Let $\mathfrak{A}$ be a τ-structure and let $\Gamma = \mathsf{D}^n_k \boldsymbol{x}\boldsymbol{i}\ \gamma(\boldsymbol{i})(\boldsymbol{x}) \in \mathbf{D}$. Then, Γ is true on $\mathfrak{A}$ iff there exist functions $f_1, \ldots, f_n : A^k \to \{0, 1\}$ such that

$$\mathfrak{A} \models \forall \boldsymbol{x}\ \gamma(f_1(\boldsymbol{x}_1), \ldots, f_n(\boldsymbol{x}_n))(\boldsymbol{x}). \tag{5}$$

Let $\Sigma^1_{n,k}(\tau)$ be the fragment of $\Sigma^1_n(\tau)$ whose predicate variables have arity $\leq k$. Particular interest will be with the fragments $\Sigma^1_{1,k}(\tau)$, that are called *k-ary, existential, second-order logic*. If k equals 1 or 2, we arrive at *monadic* and *binary*, existential, second-order logic: $\Sigma^1_{1,1}(\tau) = M\Sigma^1_1(\tau)$ and $\Sigma^1_{1,2}(\tau) = B\Sigma^1_1(\tau)$. For the semantics of first and second-order logic, we refer the reader to [6].

If Φ and Ψ are τ-sentences for which the satisfaction relation $\models$ is properly defined and for every τ-structure we have that $\mathfrak{A} \models \Phi$ iff $\mathfrak{A} \models \Psi$, then we say that they are *equivalent*.

Let $\mathbf{L}(\tau)$ be a logical language for which $\models$ is properly defined and let C be a class of (finite) τ-structures. Then C is *characterized* by $\Phi \in \mathbf{L}(\tau)$ if for every τ-structure $\mathfrak{A}$ it is the case that $\mathfrak{A} \in$ C iff $\mathfrak{A} \models \Phi$. If some of its formulae characterize the class C, then $\mathbf{L}(\tau)$ is said to characterize C as well.

Let $\mathbf{L}(\tau)$ and $\mathbf{L}'(\tau)$ be logical languages. Then, we write $\mathbf{L}(\tau) \leq \mathbf{L}'(\tau)$ to denote that for every formula $\Phi \in \mathbf{L}(\tau)$ there is an equivalent $\Psi \in \mathbf{L}'(\tau)$. We write $\mathbf{L}(\tau) = \mathbf{L}'(\tau)$, if $\mathbf{L}(\tau) \leq \mathbf{L}'(\tau)$ and $\mathbf{L}(\tau) \geq \mathbf{L}'(\tau)$. If $\mathbf{L}(\tau) \leq \mathbf{L}'(\tau)$ and there is one class characterizable in $\mathbf{L}'(\tau)$ that is not characterizable in $\mathbf{L}(\tau)$, we write $\mathbf{L}(\tau) < \mathbf{L}'(\tau)$.

By means of a game-theoretical argument we show that $\mathbf{D}$ cannot characterize the class of structures with a universe of even cardinality, Even. The latter class, however, is definable by a Henkin quantifier (with unrestricted variables).

Proposition 1. *There exists a first-order formula ϕ, such that $\mathsf{H}^2_1\ \phi$ characterizes* Even.

Proof. Recall that a finite structure $\mathfrak{A}$ has a universe A with even cardinality iff there exists a function $f : A \to A$ such that for every $a \in A$, $f(f(a)) = a$ and $f(a) \neq a$. The latter condition is expressed by the following formula: $\mathsf{H}^2_1 x_1 x_2 y_1 y_2\ \phi(x_1, x_2, y_1, y_2)$, where $\phi(x_1, x_2, y_1, y_2)$=$(x_1 = x_2 \to y_1 = y_2) \wedge (y_1 = x_2 \to y_2 = x_1) \wedge (x_1 \neq y_1)$. □

3 A characterization of $\mathbf{D}_k$

In this section we give a characterization of $\mathbf{D}_k(\tau)$ as a fragment of $\Sigma^1_{1,k}(\tau)$. First we lay down a translation result. To this end, let $\Gamma = \mathsf{D}^n_k\ \gamma$ be a $\mathbf{D}_k(\tau)$-formula, where

$$\Gamma = \mathsf{D}^n_k\ \gamma = \begin{pmatrix} \forall x_{11} \ldots \forall x_{1n} \bigvee i_1 \\ \vdots \quad \ddots \quad \vdots \quad \vdots \\ \forall x_{k1} \ldots \forall x_{kn} \bigvee i_k \end{pmatrix} \gamma\ . \tag{6}$$

Define the translation of Γ into $\Sigma^1_{1,k}(\tau)$, written $T(\Gamma)$, as follows

$$\exists X_1 \ldots \exists X_n \forall \boldsymbol{x} \begin{bmatrix} X_1(\boldsymbol{x}_1) \wedge \ \ldots \ \wedge \ \ X_n(\boldsymbol{x}_n) \to \gamma(1,\ldots,1)(\boldsymbol{x}) \\ \vdots \\ \neg X_1(\boldsymbol{x}_1) \wedge \ \ldots \ \wedge \neg X_n(\boldsymbol{x}_n) \to \gamma(0,\ldots,0)(\boldsymbol{x}) \end{bmatrix} , \tag{7}$$

where the X_i are k-ary predicate variables. The square brackets enclosing the implications should be read as their conjunction and reflects the matrix-style of presenting γ. The block of implications is referred to as *γ's explication*. The translation hinges on the insight that every function $f : A^k \to \{0,1\}$ can be mimicked by the set $X = \{a \in A^k \mid f(a) = 1\}$.

Proposition 2. *For every sentence $\Gamma \in \mathbf{D}_k(\tau)$, Γ and $T(\Gamma)$ are equivalent.*

We proceed by giving a characterization of $\mathbf{D}_k$ as a fragment of $\Sigma^1_{1,k}$.

Definition 1. *Let Φ be a second-order τ-formula. Call Φ* sober *if for every predicate variable X in Φ it is the case that (i) X is free in Φ and (ii) $X(\boldsymbol{x})$ occurring in Φ implies that all variables in $\boldsymbol{x}$ are free in Φ. Let $\Sigma^1_{1,k}\heartsuit(\tau)$ be the fragment of $\Sigma^1_{1,k}(\tau)$, containing all formulae of the form*

$$\exists X_1 \ldots \exists X_m \forall x_1 \ldots \forall x_n \ \Phi \ , \tag{8}$$

where Φ is sober. Put $\Sigma^1_1\heartsuit(\tau) = \bigcup_k \Sigma^1_{1,k}\heartsuit(\tau)$.

So any sober formula is a second-order formula, but only in virtue of the fact that it contains predicate variables. If Φ is a sober formula occurring in a $\Sigma^1_{1,k}\heartsuit(\tau)$-formula as in (8), then no variable argument to a predicate variable is existentially quantified. In Section 5 we see that the slightest extension in this respects results in a logic that enjoys the expressive power of full NP.

As an example, consider the $\Sigma^1_1\heartsuit$-formula $\exists X_1 \exists X_2 \exists X_3 \forall x_1 \forall x_2 \ (\Phi \wedge \Phi')$ that characterizes 3-Colorability, where $(\Phi \wedge \Phi')$ is a sober formula:

$$\Phi = \left(\bigvee_{i \in \{1,2,3\}} X_i(x_1) \right) \wedge \left(\bigwedge_{i \in \{1,2,3\}} \bigwedge_{j \in \{1,2,3\}-\{i\}} \neg(X_i(x_1) \wedge X_j(x_1)) \right) \tag{9}$$

$$\Phi' = \left(\bigwedge_{i \in \{1,2,3\}} (X_i(x_1) \wedge X_i(x_2) \to \neg R(x_1,x_2)) \right) . \tag{10}$$

Theorem 1. $\mathbf{D}_k(\tau) = \Sigma^1_{1,k}\heartsuit(\tau)$. *Hence,* $\mathbf{D}(\tau) = \Sigma^1_1\heartsuit(\tau)$.

Proof. The from-left-to-right direction is accounted for by the translation $T(\cdot)$. The converse direction is more involved, hinging on the proof of the claim that every sober formula is equivalent to the explication of an implicit matrix formula. $\square$

The characterization of $\mathbf{D}$ in second-order terms may speed up the finding of interesting properties it enjoys, for second-order logic happens to be more intensively studied than partially ordered connectives. Finding formulae with partially ordered connectives expressing a particular property on structures can be hard labor. Now that we have characterized $\mathbf{D}_k$, we can safely conclude that any property expressible in $\Sigma^1_{1,k}\heartsuit(\tau)$ is expressible in $\mathbf{D}_k(\tau)$ as well. A concrete – and relevant! – example of this mode of research can be found in the upcoming result.

Theorem 2. *Let $k \geq 2$ be an integer and let τ_k be a vocabulary with at least one k-ary relation symbol and $<$. Then, $\mathbf{D}_{k-1}(\tau_k) < \mathbf{D}_k(\tau_k)$.*

Proof. Ajtai [1] proved the following result: Let $k \geq 2$ and let $\tau_k = \{P, <\}$ where P is a k-ary relation symbol and $<$ is the *linear order symbol.*[4] Then, the class C_k of τ_k-structures $\mathfrak{A}$ such that $P^{\mathfrak{A}}$ has even cardinality is not characterizable in $\Sigma^1_{1,k-1}(\tau_k)$, but it is characterizable in $\Sigma^1_{1,k}(\tau_k)$.[5]

To separate $\mathbf{D}_k$ from $\mathbf{D}_{k-1}$, we show that C_k is expressible by a formula in $\mathbf{D}_k(\tau_k)$. This is a sufficient argument for our end, since $\Sigma^1_{1,k-1}\heartsuit(\tau_k)$ is a fragment of $\Sigma^1_{1,k-1}(\tau_k)$ by Theorem 1 and for this reason cannot express C_k.

Intuitively, the $\Sigma^1_{1,k}\heartsuit(\{P,<\})$-formula Υ_k that characterizes C_k over τ_k-structures lifts the linear order $<$ to a linear order ψ_k of k-tuples. With respect to this lifted linear order Υ_k expresses that there exists a subset of k-tuples of objects from the domain Q such that

1. Q is a subset of $P^{\mathfrak{A}}$
2. the ψ_k-minimal k-tuple that is in $P^{\mathfrak{A}}$ is also in Q and the ψ_k-maximal k-tuple that is in $P^{\mathfrak{A}}$ is not in Q
3. if two k-tuples are in $P^{\mathfrak{A}}$ and there is no k-tuple in between them (in the ordering constituted by ψ_k) that is in $P^{\mathfrak{A}}$, then exactly one of the k-tuples is in Q.

We omit further details. □

4 Ehrenfeucht-Fraïssé game for D

Ehrenfeucht-Fraïssé games or *model comparison games* are usually employed to prove that some property is not definable in a certain logic. These games were first introduced for first-order logic in [7, 11].

Let the *quantifier rank* of a first-order formula be its maximum number of nested quantifiers. Let m be an integer. If $\mathfrak{A}, \mathfrak{B}$ are τ-structures, $\boldsymbol{x}^{\mathfrak{A}} =$

[4] That is, on a τ_k-structure $\mathfrak{A}$, the extension of $<$ is a linear order on A.

[5] The result uses *hypergraphs*, that is, structures interpreting relation symbols of unbound arity. As a consequence, the result does not imply that $\Sigma^1_{1,2}(\tau)$ is strictly weaker than $\Sigma^1_{1,3}(\tau)$, where τ a vocabulary that contains only unary and binary predicates, cf. [5].

$\langle x_1^{\mathfrak{A}}, \ldots, x_r^{\mathfrak{A}} \rangle \in A^r$, and $\boldsymbol{x}^{\mathfrak{B}} = \langle x_1^{\mathfrak{B}}, \ldots, x_r^{\mathfrak{B}} \rangle \in B^r$, then the *m-round Ehrenfeucht-Fraïssé game on the structures* $\mathfrak{A}$ *and* $\mathfrak{B}$, denoted by

$$EF_m^{\mathbf{FO}}(\langle \mathfrak{A}, \boldsymbol{x}^{\mathfrak{A}} \rangle, \langle \mathfrak{B}, \boldsymbol{x}^{\mathfrak{B}} \rangle) \ ,$$

is an m-round game proceeding as follows: There are two players, Spoiler and Duplicator. During the ith round, Spoiler first chooses a structure $\mathfrak{A}$ (or $\mathfrak{B}$) and an element called c_i (or d_i) from the domain of the chosen structure. Duplicator replies by choosing an element d_i (or c_i) from the domain of the other structure $\mathfrak{B}$ (or $\mathfrak{A}$). Duplicator wins the play $\langle \langle c_1, d_1 \rangle, \ldots, \langle c_m, d_m \rangle \rangle$, if the relation

$$\{\langle x_i^{\mathfrak{A}}, x_i^{\mathfrak{B}} \rangle \mid 1 \le i \le r\} \cup \{\langle c_i, d_i \rangle \mid 1 \le i \le m\} \tag{11}$$

is a 'partial isomorphism' between $\mathfrak{A}$ and $\mathfrak{B}$; otherwise, Spoiler wins the play. If against any sequence of moves by Spoiler, Duplicator is able to make her moves so as to win the resulting play, we say that Duplicator has a *winning strategy in* $EF_m^{\mathbf{FO}}(\langle \mathfrak{A}, \boldsymbol{x}^{\mathfrak{A}} \rangle, \langle \mathfrak{B}, \boldsymbol{x}^{\mathfrak{B}} \rangle)$. The notion of winning strategy for Spoiler is defined analogously. By the Gale-Stewart Theorem [12], the Ehrenfeucht-Fraissé games are determined; that is, precisely one of the players has a winning strategy. The effectiveness of these games is established in the following seminal result.

Theorem 3 ([7, 11]). *For every integer m, the following are equivalent:*

- $\langle \mathfrak{A}, \boldsymbol{x}^{\mathfrak{A}} \rangle$ *and* $\langle \mathfrak{B}, \boldsymbol{x}^{\mathfrak{B}} \rangle$ *satisfy the same first-order formulae (possibly with free variables from* $\boldsymbol{x}$*) of quantifier rank* $\le m$
- *Duplicator has a winning strategy in* $EF_m^{\mathbf{FO}}(\langle \mathfrak{A}, \boldsymbol{x}^{\mathfrak{A}} \rangle, \langle \mathfrak{B}, \boldsymbol{x}^{\mathfrak{B}} \rangle)$.

Readers unfamiliar with these games may find it helpful to consult [6] and [10, 15] for a similar games for $M\Sigma_1^1$.

The notion of quantifier rank is extended to implicit matrix formulae as follows: $qr(\gamma) = \max\{qr(\gamma(\boldsymbol{i})) \mid \boldsymbol{i} \in \{0,1\}^k\}$, for γ of type $\{0,1\}^k \to \mathbf{FO}$.

The model comparison game for $\mathbf{D}$ has two phases: a *watercoloring phase* and a *first-order phase*. Let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures and let m be an integer. Then, the *m-round, watercolor* $\mathbf{D}_k^n$*-Ehrenfeucht-Fraïssé game on the structures* $\mathfrak{A}$ *and* $\mathfrak{B}$, denoted as

$$EF_m^{\mathbf{D}_k^n}(\mathfrak{A}, \mathfrak{B})$$

is an $m+1$-round game proceeding as follows: First we have the watercoloring phase. Spoiler picks up for every $1 \le i \le n$ a subset X_i from A^k. Duplicator picks up a subset B_i of B^k, for every $1 \le i \le n$. Next, Spoiler chooses a tuple $\boldsymbol{x}_i^{\mathfrak{B}} \in B^k$, for every $1 \le i \le n$, and Duplicator replies by choosing a tuple $\boldsymbol{x}_i^{\mathfrak{A}} \in A^k$. If for every $1 \le i \le n$ the selected tuples satisfy $\boldsymbol{x}_i^{\mathfrak{A}} \in A_i \quad \text{iff} \quad \boldsymbol{x}_i^{\mathfrak{B}} \in B_i$, then the game proceeds to the first-order phase as $EF_m^{\mathbf{FO}}(\langle \mathfrak{A}, \boldsymbol{x}^{\mathfrak{A}} \rangle, \langle \mathfrak{B}, \boldsymbol{x}^{\mathfrak{B}} \rangle)$; otherwise, Duplicator loses right away.

Interesting to note that in the first-order Ehrenfeucht-Fraïssé game that is started up after the watercolor phase, the actual colorings are immaterial. The watercolors fade away quickly, so to say.

Proposition 3. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures, and let k, n be integers. Let $\Gamma = \mathsf{D}_k^n\ \gamma$ be any $\mathbf{D}_k$-sentence with $qr(\gamma) \leq m$. Then, the first assertion implies the second:*

- *Duplicator has a winning strategy in $EF_m^{\mathbf{D}_k^n}(\mathfrak{A}, \mathfrak{B})$*
- *$\mathfrak{A} \models \Gamma$ implies $\mathfrak{B} \models \Gamma$.*

Hence, if the first assertion holds for arbitrary k, n, then the second assertion holds for every $\Gamma \in \mathbf{D}$, where $qr(\Gamma) \leq m$.

Proof. The game is a simple adaption of the one presented in [19]. $\square$

Fagin [10] showed that the monadic fragments of Σ_1^1 and Π_1^1 do not coincide, as the latter harbors CONNECTED but the former does not. Therefore, $M\Sigma_1^1 \neq \mathbf{M}\Pi_1^1$ and we say that Σ_1^1 *is not closed under complementation.*

Using the model comparison games for $\mathbf{D}$ we show that $\mathbf{D}$ is not closed under complementation either. This result may be interesting because $\mathbf{D} = \Sigma_1^1\heartsuit$ is a fragment of Σ_1^1 that is not bounded by the arity of the predicate variables and has a non-empty intersection with k-ary, existential, second-order logic, for arbitrary k, see Theorem 2. Clearly, these properties are not enjoyed by $M\Sigma_1^1$.

For any two τ-structures $\mathfrak{A}$ and $\mathfrak{B}$ with non-intersecting universes, let $\mathfrak{A} \dot\cup \mathfrak{B}$ denote the τ-structure with universe $A \cup B$ and $R^{\mathfrak{A}\dot\cup\mathfrak{B}} = R^{\mathfrak{A}} \cup R^{\mathfrak{B}}$, for any $R \in \tau$.

Theorem 4. $\overline{\text{2-COLORABILITY}}$ *cannot be expressed in* $\mathbf{D}$. *Hence,* $\mathbf{D}$ *is not closed under complementation.*

Proof. For the sake of contradiction, suppose $\overline{\text{2-COLORABILITY}}$ were characterizable in $\mathbf{D}$. Then, there would be a particular formula in $\mathbf{D}$ that characterizes $\overline{\text{2-COLORABILITY}}$, say Γ. This formula Γ would be have a partially ordered connective with dimensions k, n prefixing a implicit matrix τ-formula of quantifier rank m. Now let us consider two structures $\mathfrak{A}$ and $\mathfrak{B}$ such that (i) $\mathfrak{A}$ is not 2-colorable and $\mathfrak{B}$ is 2-colorable and (ii) Duplicator has a winning strategy in $EF_m^{\mathbf{D}_k^n}(\mathfrak{A}, \mathfrak{B})$. Since Γ is supposed to characterize $\overline{\text{2-COLORABILITY}}$, we derive from (i) that $\mathfrak{A} \models \Gamma$ and $\mathfrak{B} \not\models \Gamma$. But from (ii) and $\mathfrak{A} \models \Gamma$ it follows by Proposition 3, that $\mathfrak{B} \models \Gamma$. Contradiction. Hence, no formula Γ exists in $\mathbf{D}$, expressing $\overline{\text{2-COLORABILITY}}$.

It remains to be shown that for arbitrary m, k, n, there exist graphs $\mathfrak{A}$ and $\mathfrak{B}$ meeting (i) and (ii). To this end, fix integers m, k, n and consider the graphs $\mathfrak{C}$ and $\mathfrak{D}$, where

$$
\begin{aligned}
C &= \{c_1, \ldots, c_N\} \\
R^{\mathfrak{C}} &= \text{the symmetric closure of } \{\langle c_i, c_{i+1}\rangle \mid 1 \leq i \leq N-1\} \cup \{\langle c_N, c_1\rangle\} \\
D &= \{d_1, \ldots, d_{N+1}\} \\
R^{\mathfrak{D}} &= \text{the symmetric closure of } \{\langle d_i, d_{i+1}\rangle \mid 1 \leq i \leq N\} \cup \{\langle d_{N+1}, d_1\rangle\}
\end{aligned}
$$

and $N = 2^{m+k\cdot n}$. So $\mathfrak{C}$ and $\mathfrak{D}$ are cycles of even and odd length, respectively. A cycle is 2-colorable iff it is of even length, hence $\mathfrak{D}$ is not 2-colorable whereas $\mathfrak{C}$ is. Obviously, the structure $\mathfrak{C} \dot\cup \mathfrak{D}$ is not 2-colorable either.

Let us proceed to showing that Duplicator has a winning strategy in $EF_m^{\mathbf{D}_k^n}(\mathfrak{C} \dot\cup \mathfrak{D}, \mathfrak{C})$. Suppose Spoiler selects the set $X_i \subseteq (C \cup D)^k$, for every $1 \leq i \leq n$. Let Duplicator respond with X_i restricted to $\mathfrak{C}$ solely, that is, $Y_i = X_i \cap C^2$, for every $1 \leq i \leq n$. Suppose Spoiler selects the tuple $\boldsymbol{x}_i^{\mathfrak{C}} \in C^k$, for every $1 \leq i \leq n$. Let Duplicator respond by simply copying these tuples on $(C \cup D)^k$, that is, setting $\boldsymbol{x}_i^{\mathfrak{C} \dot\cup \mathfrak{D}} = \boldsymbol{x}_i^{\mathfrak{C}}$. The game advances to the first-order phase, since obviously $\boldsymbol{b}^i \in X_i$ iff $\boldsymbol{b}^i \in Y_i$. A standard argument suffices to see that Duplicator has a winning strategy in

$$EF_m^{\mathbf{FO}}(\langle \mathfrak{C} \dot\cup \mathfrak{D}, \boldsymbol{x}_1^{\mathfrak{C} \dot\cup \mathfrak{D}}, \ldots, \boldsymbol{x}_n^{\mathfrak{C} \dot\cup \mathfrak{D}} \rangle, \langle \mathfrak{C}, \boldsymbol{x}_1^{\mathfrak{C}}, \ldots, \boldsymbol{x}_n^{\mathfrak{C}} \rangle) \ ,$$

compare [6, pg. 23].

In the Introduction we recalled that Blass and Gurevich showed that $\mathbf{D}$ can characterize the class of 3-colorable graphs. In the same way it is capable of characterizing 2-Colorability. We just showed that the complement of this class is not expressible in $\mathbf{D}$. Therefore, $\mathbf{D}$ is not closed under complementation. □

Since $\mathfrak{C}$'s universe has even cardinality but $\mathfrak{D}$'s has not, we conclude that also the class Even is not characterizable in $\mathbf{D}$. By contrast, in Proposition 1 we showed that this class is characterizable by a formula of the form $\mathsf{H}_1^2\ \phi$. So already the simplest Henkin quantifier, not definable in first-order logic, cannot be defined in $\mathbf{D}$. Since Even is obviously characterizable in binary Σ_1^1, $\mathbf{D} < \Sigma_1^1$.

5 Revisiting $\Sigma_1^1\heartsuit$

We mapped out some finite model theory for $\mathbf{D}$ and observed that it is closed under complementation but not bounded by an arity constraint. We saw that $\mathbf{D}$ comprises a fragment of Σ_1^1 whose formulae do not allow for a single existential variable being an argument of a predicate variable. Amusingly, this boundary is rather sharp: already the slightest extension yields a logic coinciding with Σ_1^1. Let us write $\Sigma_1^1\clubsuit$ for the fragment of Σ_1^1 that has formulae of the form

$$\exists X_1 \ldots \exists X_m \mathsf{Q}_1 x_1 \ldots \mathsf{Q}_n x_n\ \Phi \tag{12}$$

where Φ is sober as before and for at most one $i \in \{1, \ldots, n\}$, we have that $\mathsf{Q}_i = \exists$; all other quantifiers are universal quantifiers. Using a result of Krynicki's [16] it is not so hard to see that $\Sigma_1^1\clubsuit = \mathrm{NP}$ on finite structures. Krynicki showed, namely, that first-order logic prefixed by the quantifier below (with unbound k) coincides with full Σ_1^1:

$$\begin{pmatrix} \forall x_{11} \ldots \forall x_{1k} \bigvee i \\ \forall x_{21} \ldots \forall x_{2k}\ \exists y \end{pmatrix} \ . \tag{13}$$

The semantics of (13) are readily defined in view of the semantics of (1) and (4), involving one function variable of type $A^k \to \{0,1\}$ and one function variable of type $A^k \to A$. The former function variable can be mimicked by a k-ary predicate variable as in the translation $T(\cdot)$. The latter k-ary function variable can be mimicked by a $k+1$-ary predicate variable along the obvious path, be it at the cost of introducing one existential quantifier.

References

1. M. Ajtai. Σ_1^1-formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.
2. M. Ajtai and R. Fagin. Reachability is harder for directed than for undirected graphs. *Journal of Symbolic Logic*, 55:113–150, 1990.
3. M. Ajtai, R. Fagin, and L. Stockmeyer. The closure of monadic NP. *Journal of Computer and System Sciences*, 60(3):660–716, 2000.
4. A. Blass and Y. Gurevich. Henkin quantifiers and complete problems. *Annals of Pure and Applied Logic*, 32:1–16, 1986.
5. A. Durand, C. Lautemann, and T. Schwentick. Subclasses of Binary-NP. *Journal of Logic and Computation*, 8(2):189–207, 1998.
6. H.-D. Ebbinghaus and J. Flum. *Finite Model Theory.* Springer-Verlag, Berlin, 1999.
7. A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141, 1961.
8. H.B. Enderton. Finite partially ordered quantifiers. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 16:393–397, 1970.
9. R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. *Complexity of Computation, SIAM-AMS Proceedings*, 7, 1974.
10. R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.
11. R. Fraissé. Sur quelques classifications des systèmes de relations. Publications Scientifiques, Série A, 35–182 1, Université d'Alger, 1954.
12. D. Gale and F. Stewart. Infinite games with perfect information. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II*, volume 28 of *Annals of Mathematics Studies*, pages 245–266. Princeton University Press, Princeton NJ, 1953.
13. G. Gottlob. Relativized logspace and generalized quantifiers over finite ordered structures. *The Journal of Symbolic Logic*, 62(2):545–574, 1997.
14. L. Hella and G. Sandu. Partially ordered connectives and finite graphs. In M. Krynicki, M. Mostowski, and L.W. Szczerba, editors, *Quantifiers: Logics, Models and Computation*, volume II of *Synthese library: studies in epistemology, logic, methodology, and philosophy of science*, pages 79–88. Kluwer Academic Publishers, Dordrecht, 1995.
15. N. Immerman. *Descriptive Complexity.* Graduate texts in computer science. Springer, New York, 1999.
16. M. Krynicki. Hierarchies of finite partially ordered connectives and quantifiers. *Mathematical Logic Quarterly*, 39:287–294, 1993.
17. M. Krynicki and M. Mostowski. Henkin quantifiers. In M. Krynicki, M. Mostowski, and L.W. Szczerba, editors, *Quantifiers: Logics, Models and Computation*, volume I of *Synthese library: studies in epistemology, logic, methodology, and philosophy of science*, pages 193–262. Kluwer Academic Publishers, The Netherlands, 1995.
18. G. Sandu. The logic of informational independence and finite models. *Logic Journal of the IGPL*, 5(1):79–95, 1997.
19. G. Sandu and J. Väänänen. Partially ordered connectives. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 38:361–372, 1992.
20. G. Turán. On the definability of properties of finite graphs. *Discrete Mathematics*, 49:291–302, 1984.
21. W. Walkoe. Finite partially-ordered quantification. *Journal of Symbolic Logic*, 35:535–555, 1970.