Differential Games, Locality, and Model Checking for FO Logic of Graphs

Jakub Gajarský ✉
University of Warsaw, Poland

Maximilian Gorsky ✉
TU Berlin, Germany

Stephan Kreutzer ✉
TU Berlin, Germany

Abstract

We introduce differential games for FO logic of graphs, a variant of Ehrenfeucht-Fraïssé games in which the game is played on only one graph and the moves of both players are restricted. We prove that these games are strong enough to capture essential information about graphs from graph classes which are interpretable in nowhere dense graph classes. This, together with the newly introduced notion of differential locality and the fact that the restriction of possible moves by the players makes it easy to decide the winner of the game in some cases, leads to a new approach to the FO model checking problem which can be used on various graph classes interpretable in classes of sparse graphs.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Theory of computation → Fixed parameter tractability

Keywords and phrases FO model checking, locality, Gaifman's theorem, EF games

Digital Object Identifier 10.4230/LIPIcs.CSL.2022.22

Related Version *Full Version*: `https://arxiv.org/pdf/2007.11345.pdf`

Funding *Jakub Gajarský*: The research reported in this paper is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC consolidator grant LIPA, agreement No. 683080).
Maximilian Gorsky: The research reported in this paper has been supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC consolidator grant DISTRUCT, agreement No. 648527).
Stephan Kreutzer: The research reported in this paper has been supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC consolidator grant DISTRUCT, agreement No. 648527).

1 Introduction

The first-order (FO) model checking problem asks, given a graph G and a sentence φ as input, whether $G \models \varphi$. It is known that this problem is PSPACE-complete in general [26, 27], but one can obtain efficient parameterised algorithms on many structurally restricted classes of graphs. Here by efficient parameterised algorithms we mean algorithms with runtime $f(|\varphi|) \cdot n^{O(1)}$; these are known as fpt algorithms.

There has been a long line of research studying this problem on sparse graphs and the existence of fpt algorithms was established for graphs of bounded degree [25], graphs with locally bounded treewidth [13], graphs with a locally excluded minor [6], bounded expansion graph classes [7] and nowhere dense graph classes [19]. The positive results on non-sparse graphs fall into two categories. The first category are formed by somewhat isolated results such as [18, 15, 21, 10] and the recent important and general result of [1]. The second category are positive results about graph classes which can be obtained from sparse graph classes by means of interpretations [16, 17] (although [10] can also be put into this category).

One of the reasons why the research into the FO model checking has been so successful is that Gaifman's theorem [14] – an important result which essentially states that FO logic is local – is particularly useful in the case of sparse graphs. Informally, Gaifman's theorem allows us to reduce the problem of determining whether a given FO formula φ holds on a given graph G to the problem of evaluating a formula $\psi(x)$ in the r-neighbourhood of each vertex of G. In case G is a graph in which each vertex has a simple neighbourhood, one can evaluate φ on G efficiently. This idea leads to efficient algorithms for evaluating FO formulas on classes of graphs of bounded degree, planar graphs, and graphs with locally bounded treewidth.

One shortcoming of using Gaifman's theorem for evaluating FO formulas is that if a graph has an (almost) universal vertex, then for $r \geq 2$ the r-neighbourhood of any vertex is (almost) the whole graph, and therefore evaluating formulas locally on the r-neighbourhoods is essentially the same as evaluating them on the whole graph. Even worse, on complements of bounded degree graphs, it holds for every vertex v that almost the whole graph is in the 1-neighbourhood of v. In such cases, one cannot use the locality-based approach directly, but has to complement the input graph G to get the graph $\bar{G}$ first, use the locality of $\bar{G}$, and then translate the results back to G. In many cases when dealing with non-sparse graphs, there seems to be no good way of using Gaifman's theorem at all, and either one uses a notion of locality tailor-made to the given situation (such as in [15] or [1]) or does not use locality at all (for example the dynamic programming algorithm for FO (and even MSO) logic on graph classes of bounded treewidth).

In this paper we initiate a relativised approach to FO model checking, which is aimed to work on graph classes interpretable in nowhere dense graph classes and which avoids some of the issues mentioned above. Instead of focusing on the absolute question "What is the r-local q-type of a vertex v?" (which is essentially what we do when we apply Gaifman's theorem to obtain algorithmic results), our approach is based on the relative question that, for a pair u, v of vertices, asks "Is the q-type of u and v the same?". It is not difficult to show that being able to answer this question efficiently leads to an efficient algorithm for the FO model checking problem.

The advantage of the relativised approach stems from the fact that one may be able to determine whether q-types of u and v are the same or not without actually determining their q-types themselves – for example if u and v are twins, then they necessarily will have the same q-type, and if one of them has two neighbours and the other has three neighbours, then they cannot have the same q-type (for $q \geq 3$). Moreover, as the examples just given suggest, it is true that if the q-types of u and v are different, then it is possible to exhibit some difference between them in $D(u, v)$ – the symmetric difference of neighbourhoods of u and v. To make this precise, we introduce the *differential game*, a newly defined version of Ehrenfeucht-Fraïssé game, which is played between two vertices u, v of a graph G (the whole game is played on one graph only) and in which the moves of the players are restricted – the first move takes place in $D(u, v)$, the second move in $D(u, v) \cup D(u_1, v_1)$ (here u_1 and v_1 are the vertices played in the first move) and so on. The potential algorithmic advantage offered by our relativized approach comes from the fact that the set $D(u, v)$ can be significantly smaller than sets $N(u)$ and $N(v)$ – for example in complements of graphs of degree at most d, the set $D(u, v)$ has size at most $2d + 2$ while $N(u)$ and $N(v)$ are large.

Our contributions can be briefly summarised as follows:

1. We show that FO model checking can be reduced to deciding whether two vertices u, v of a graph G have the same q-type, i.e. whether $u \equiv_q v$. Note that this is not entirely

trivial – even if we have access to the q-equivalence relation on $V(G)$, it is not clear which equivalence class corresponds to which q-type.

2. We introduce *differential games* which are aimed at distinguishing vertices of different q-types. We prove that for every q there exists r such that the relation $u \cong_r^D v$ defined by "Duplicator wins the r-round differential game between u and v" suitably approximates $\equiv_q$ on graph classes interpretable in nowhere dense graph classes. This leads to the following theorem:

   ▶ **Theorem 6.3.** *Let $\mathcal{C}$ be a class of labelled graphs interpretable in a nowhere dense class of graphs such that we can decide the winner of the r-round differential game in fpt runtime with respect to the parameter r. Then the FO model checking problem is solvable in fpt runtime on $\mathcal{C}$.*

3. As a simple application of our methods, we reprove the result that FO model checking is in FPT for graph classes interpretable in classes of graphs of bounded degree. The important aspect of our new proof is that our algorithm does not involve computing any decomposition of the input graph (in particular, it does not rely on computing a sparse pre-image of the input graph).
4. We define the *differential r-neighborhood* $DN_r(u, v)$ of two vertices, which is an extension of $D(u, v)$ to a "larger radius" and which has the property that the whole r-round differential game played between u and v is played in $DN_r(u, v)$. We use differential neighborhoods to define *differentially simple* graph classes, which are graph classes on which it is possible to decide the winner of the r-round differential game efficiently.

Our results do not immediately lead to new strong FO model checking results – at this point our main contribution is mainly conceptual. We believe that the tools and ideas presented here may lead to new insights and algorithms for graph classes interpretable in classes of sparse graphs. In particular, the idea of performing model checking directly on the input graph, without computing any decomposition or sparsification, seems interesting and probably deserves further attention.

Organization. We give an overview of our ideas in the next section. Section 4 is devoted to relativized model checking, and Sections 5 and 6 are devoted to differential games. Applications are discussed in sections 7 and 8.

2 Overview of our approach

As mentioned in the introduction, our relativised approach to FO model checking is based on determining whether two vertices u, v of G *differ* from each other – i.e. whether $u \not\equiv_q v$, which is the case whenever there is a formula $\psi(x)$ of quantifier rank q such that $G \models \psi(u)$ but $G \not\models \psi(v)$. In Section 4, we show that if we can solve this problem efficiently, then we can solve the FO model checking problem efficiently as well. To be more precise, we show that if we can efficiently compute a relation $\sim_q$ on $V(G)$ such that the transitive closure of $\sim_q$ refines $\equiv_q$ and does not have too many classes, then we can construct an *evaluation tree* of size bounded in terms of q. This evaluation tree then allows us to determine whether $G \models \varphi$ for every sentence φ in prenex normal form with q quantifiers.

From this point on, we focus on efficiently determining whether two vertices u, v differ. Our approach relies on the intuition that if they differ, then there should be a way to exhibit the difference through $D(u, v)$. To capture this intuition formally, we introduce *semi-differential games* in Section 5, which are a variant of the well-known Ehrenfeucht-Fraïssé

(EF) games. The semi-differential game is played in one graph only and if we are given two vertices a_0, b_0 as a starting position, we play the game with the Spoiler's moves being restricted in the following way: In his first move the Spoiler plays a vertex in $D(a_0, b_0)$ and declares the picked vertex to be either a_1 or b_1. The Duplicator picks her reply anywhere in G and the picked vertex becomes b_1 or a_1 – the "opposite" of the Spoiler's choice. More generally, in the i-th move, after the vertices $a_1, \ldots, a_{i-1}$ and $b_1, \ldots, b_{i-1}$ have been played, the Spoiler picks $j \in \{0, \ldots, i-1\}$ and a vertex in $D(a_j, b_j)$ and calls it a_i or b_i. Again, the Duplicator replies by picking a vertex anywhere in G. The winner of the game is decided as in the usual EF game, by comparing the graphs induced by $(a_0, a_1, \ldots, a_m)$ and $(b_0, b_1, \ldots, b_m)$.

As our first crucial result, we show (Lemma 5.2) that there exists a function $l : \mathbb{N} \to \mathbb{N}$ such that if the Spoiler wins the standard r-round Ehrenfeucht-Fraïssé game on a graph G starting from a_0 and b_0, then he wins the $l(r)$-round semi-differential game starting from the same position. This is the aforementioned formalization of the fact that the difference between two vertices can be exhibited through $D(u, v)$. If we denote by $u \cong_r^{SD} v$ the relation "Duplicator wins the r-round semi-differential game starting from u and v", then the above result also tells us that $\cong_{l(q)}^{SD}$ refines $\equiv_q$, and so we can use $\cong_{l(q)}^{SD}$ as the relation $\sim_q$ from the first paragraph of this section (it is not difficult to show that the transitive closure of $\cong_{l(q)}^{SD}$ has bounded number of classes, see Lemma 5.3).

While the relation $\cong_{l(q)}^{SD}$ could serve as a suitable refinement $\sim_q$ of $\equiv_q$, the fact that the Duplicator's moves are not restricted in any way makes it hard to use it algorithmically. To alleviate this, we introduce *differential games* in Section 6 in which the Duplicator's moves are restricted to $D(a_j, b_j)$ as well. For every graph G and every r we define the relation on $V(G)$ by setting $u \cong_r^D v$ if and only if Duplicator wins the r-round differential game starting from u and v. Note that $\cong_r^D$ is a subset of $\cong_r^{SD}$, because whenever Duplicator wins the differential game, she also wins the semi-differential game. This means that the relation $\cong_{l(q)}^D$ also refines $\equiv_q$, as required. Unfortunately, the transitive closure of relation $\cong_r^D$ does not have bounded number of classes, in terms of r, in general. An example of this are ladders – bipartite graphs on vertex set $\{v_1, \ldots, v_{2n}\}$ where the two parts are formed by even and odd numbered vertices and in which there is an edge between v_i and v_j where i is odd and j is even if $i < j$. The reason is that each side in a ladder contains n vertices with nested neighbourhoods, and for any u, v with $N(u) \subsetneq N(v)$ the Spoiler wins the 1-round differential game. There are, however, very rich classes of graphs which exclude arbitrarily large ladders in a very strong sense – such classes of graphs (and more general structures) are known in model theory as *stable* classes of graphs (structures). A prominent example are nowhere dense graph classes introduced by Nešetřil and Ossona de Mendez [22, 23]. On these classes of graphs we can show that the graph of the relation $\cong_m^D$ has a bounded number of connected components. Moreover, we can also show this for graph classes interpretable in nowhere dense graph classes (Theorem 6.2). This leads us to the conclusion that if we can decide the winner of the r-round differential game on any class of graphs which is interpretable in a nowhere dense graph class, then we can perform FO model checking efficiently.

We now turn our attention to classes of graphs on which the winner of the differential game can be found efficiently. First we describe the main idea behind the notion of *differential r-neighbourhood*. In its simplest form it is defined as follows: $DN_1(u, v)$ is just $D(u, v)$ and for any $i > 1$, the differential i-neighbourhood $DN_i(u, v)$ is $DN_{i-1}(u, v)$ together with the union of all $D(a, b)$, with $a, b \in DN_{i-1}(u, v)$. It is easy to see that the whole r-round differential game starting from u, v takes place in $DN_r(u, v)$. If the subgraph of G induced by $DN_r(u, v)$ is simple (say, it has small treewidth), we can decide the winner of differential game efficiently – one can just write a formula $\xi_r(x, y)$ saying "Duplicator wins the r-round differential game between x and y" and evaluate it on $DN_r(u, v)$ using Courcelle's theorem in our example. This is the essential idea behind *differentially simple* graph classes, in which for all graphs G and all pairs of vertices u, v it holds that $DN_r(u, v)$ induces a subgraph of G which is simple (comes from a class of graphs with efficient FO model checking algorithm). It is easily seen that the "usual suspects" – graphs of bounded degree and of locally bounded treewidth are differentially simple, and so are the classes of complements of these graphs.

We remark that the actual definition of differential neighbourhoods given in Section 8 is slightly more involved and uses colourings. On one hand this makes the notion more general (interpretations of classes of locally bounded treewidth are differentially simple, see Lemma 8.4), on the other hand computing the required colourings may be difficult – we were unable to find a polynomial-time algorithm which computes them. This is similar to the results of [17] and [1], in which the existence of a model checking algorithm is proven, provided that a suitable decomposition of the input graph is given.

3 Preliminaries

We use standard notation from graph theory. All graphs in this paper are finite, undirected, simple, and without loops. The depth of a rooted tree T is the largest number of edges on any leaf-to-root path in T and we say that a node p is at depth i in T if the distance of p from the root of T is i.

By $A\Delta B$ we denote the symmetric difference of two sets A and B defined by $A\Delta B = (A \setminus B) \cup (B \setminus A)$.

3.1 Logic

We assume familiarity with FO logic. We refer to [8] or any standard logic textbook for precise definitions. Since in the paper we only work with finite, simple, undirected graphs, to simplify the exposition we define the notions from logic and model theory for the vocabulary $\sigma = \{E, \{L_a\}_{a \in Lab}\}$ of labelled graphs. Here E is a binary relation symbol, Lab is a finite set of labels and each L_a is a unary predicate symbol.

We say that two graphs G and H are m-equivalent, denoted by $G \equiv_m H$, if they satisfy the same FO sentences of quantifier rank m. It is well known that for every m the relation $\equiv_m$ is an equivalence with finitely many classes.

The FO q-type of a tuple of vertices $\bar{a} = (a_1, \ldots, a_k) \in V(G)^k$, for a given (labelled) graph G, is defined as the set of formulas $\mathrm{tp}_q^G(\bar{a}) \coloneqq \{\psi(x_1, \ldots, x_k) \in \mathrm{FO}[\sigma] \mid G \models \psi(a_1, \ldots, a_k)$ and ψ has quantifier rank $q\}$, where $\sigma = \{E\}$, or $\sigma = \{E, \{L_a\}_{a \in Lab}\}$, if G is labelled with elements of Lab.

Using the notion of q-types, we can more generally define, for the tuples $\bar{v} \coloneqq (v_1, \ldots, v_k)$ and $\bar{u} \coloneqq (u_1, \ldots, u_k)$, consisting of vertices from G, and respectively from H, that $(G, \bar{v}) \equiv_q^k (H, \bar{u})$ if and only if $\mathrm{tp}_q^G(\bar{v}) = \mathrm{tp}_q^H(\bar{u})$. We will mostly be interested in the case when $G = H$; whenever we write $\bar{v} \equiv_q^k \bar{u}$, it is understood that $\bar{v}$ and $\bar{u}$ come from the same graph G, which is clear from the context and should we want to refer to the relation itself and need to note the graph it is based upon, we will add the graph as an index, as in $\equiv_q^{k,G}$. Note that up to equivalence there exist only a finite number of formulas with a given quantifier rank and number of free variables. Therefore there also only exist a finite number of q-types for any given number of free variables. Thus the graph of the relation $\equiv_q^{k,G}$ has a number of components (cliques) bounded by a number depending only on q and k.

For a graph G and a tuple $\bar{v} = (v_1, \ldots, v_k)$ of vertices of G, we define the relation $\equiv_q^{\bar{v}}$ on $V(G)$ by setting $u \equiv_q^{\bar{v}} w$ if and only if $(v_1, \ldots, v_k, u) \equiv_q (v_1, \ldots, v_k, w)$.

3.2 Games

Let G and H be two graphs, and $m \in \mathbb{N}$. The m-round *Ehrenfeucht-Fraïssé game* [11, 12, 9] (or EF game for short), denoted by $\mathcal{G}_m(G, H)$, is played by two players called the *Spoiler* and the *Duplicator*. Each player has to make m moves in the course of play. The players take turns, with the Spoiler going first in each round. In his i-th move, the Spoiler first selects a graph, G or H, and a vertex in this graph. If the Spoiler chooses v_i in G then the Duplicator in her i-th move must choose an element u_i in H. If the Spoiler chooses u_i in H then Duplicator in her i-th move must choose an element v_i in G. The Duplicator wins if $\iota(v_i) = u_i$ is a label preserving isomorphism from $G[\{v_1, \dots, v_m\}]$ to $H[\{u_1, \dots, u_m\}]$. Otherwise the Spoiler wins. We say that a player has a *winning strategy*, or in short that they *win* $\mathcal{G}_m(G, H)$, if it is possible for them to win each play regardless of the choices made by their opponent. We denote the fact that the Duplicator wins the m-round EF game between graphs G and H by $G \cong_m H$. The relation $\cong_m$ is an equivalence with finitely many classes for every m. The following theorem connects EF games and m-equivalence.

▶ **Theorem 3.1** (Corollary 2.2.9 in [8]). *Let G and H be graphs and $m \in \mathbb{N}$. Then $G \equiv_m H$ if and only if $G \cong_m H$.*

A *position* in $\mathcal{G}_m(G, H)$ is a tuple $((v_1, \dots, v_k), (u_1, \dots, u_{k'}))$, where each v_i is from $V(G)$, each u_i is from $V(H)$, and it holds that $k, k' \leq m$ and $|k - k'| \leq 1$. If $|k - k'| = 0$, then it is the Spoiler's move, otherwise it is the Duplicator's move.

Let G and H be graphs, $(v_1, \dots, v_k)$ a tuple of vertices of G and $(u_1, \dots, u_k)$ a tuple of vertices of H. For every m we can play the m-round EF game between $(v_1, \dots, v_k)$ and $(u_1, \dots, u_k)$, denoted as $\mathcal{G}_m((G, v_1, \dots, v_k), (H, u_1, \dots, u_k))$, by considering the $(k+m)$-round EF game between G and H in which the position $((v_1, \dots, v_k), (u_1, \dots, u_k))$ has been reached and starting the play from this position. If the Duplicator wins the m-round game between $(v_1, \dots, v_k)$ and $(u_1, \dots, u_k)$, we denote this by $(G, v_1, \dots, v_k) \cong^k_m (H, u_1, \dots, u_k)$. The following more general version of Theorem 3.1 connects relations $\equiv^k_m$ and $\cong^k_m$.

▶ **Theorem 3.2** (Theorem 2.2.8 in [8]). *Let G and H be graphs, $(v_1, \dots, v_k)$ a tuple of vertices of G, $(u_1, \dots, u_k)$ a tuple of vertices of H, and m a non-negative integer. Then $(G, v_1, \dots, v_k) \equiv^k_m (H, u_1, \dots, u_k)$ if and only if $(G, v_1, \dots, v_k) \cong^k_m (H, u_1, \dots, u_k)$.*

Again we will be mostly interested in the case when $G = H$; whenever we write $(v_1, \dots, v_k) \cong^k_m (u_1, \dots, u_k)$ it is understood that $(v_1, \dots, v_k)$ and $(u_1, \dots, u_k)$ come from the same graph G which is clear from the context. When comparing two concrete tuples, we will write $\cong_m$ instead of $\cong^k_m$, since k can be inferred from the context and we will apply the same rationale to $\equiv^k_m$ as well.

3.3 Interpretations

Let $\psi(x, y)$ be an FO formula with two free variables over the language of (possibly labelled) graphs such that for any graph and any u, v it holds that $G \models \psi(u, v) \Leftrightarrow G \models \psi(v, u)$ and $G \not\models \psi(u, u)$, i.e. the relation on $V(G)$ defined by the formula is symmetric and irreflexive. From now on we will assume that formulas with two free variables are symmetric and irreflexive (which can easily be enforced). Given a graph G, the formula $\psi(x, y)$ maps G to a graph $H = I_\psi(G)$ defined by $V(H) = V(G)$ and $E(H) = \{\{u, v\} \mid G \models \psi(u, v)\}$. We then say that the graph H is *interpreted* in G.

In case G is labelled, H inherits labels from G. This way, whenever we need graph H to be labelled, it is enough to consider appropriately labelled G. In case we do not need certain labels from G in H we can simply ignore or drop them when necessary.

The notion of interpretation can be extended to graph classes as well. To a graph class $\mathcal{C}$ the formula $\psi(x,y)$ assigns the graph class $\mathcal{D} = I_\psi(\mathcal{C}) = \{H \mid H = I_\psi(G), G \in \mathcal{C}\}$. We say that a graph class $\mathcal{D}$ is *interpretable* in a graph class $\mathcal{C}$ if there exists formula $\psi(x,y)$ such that $\mathcal{D} \subseteq I_\psi(\mathcal{C})$. Note that we do not require $\mathcal{D} = I_\psi(\mathcal{C})$, as we just want every graph from $\mathcal{D}$ to have a preimage in $\mathcal{C}$.

3.4 Gaifman's theorem

An FO formula $\varphi(x_1,\ldots,x_l)$ is r-*local* if for every graph G and all $v_1,\ldots,v_l \in V(G)$ it holds that $G \models \varphi(v_1,\ldots,v_l) \iff \bigcup_{1\le i\le l} N_r^G(v_i) \models \varphi(v_1,\ldots,v_l)$, where $N_r^G(v)$ is the subgraph of G induced by v and all vertices of distance at most r from v.

▶ **Theorem 3.3** (Gaifman's theorem, [14]). *Every first-order formula with free variables $x_1,\ldots,x_l$ is equivalent to a Boolean combination of the following:*

- *Local formulas $\phi^{(r)}(x_1,\ldots,x_l)$ around $x_1,\ldots,x_l$, and*
- *Basic local sentences, i.e. sentences of the form*

$$\exists x_1 \ldots \exists x_k \left(\bigwedge_{1\le i<j\le k} dist(x_i,x_j) > 2r \wedge \bigwedge_{1\le i\le k} \phi^{(r)}(x_i) \right)$$

where each $\phi^{(r)}(x_i)$ is a r-local formula

We will need the following simple corollary of Gaifman's theorem, in which we denote by $tp_q^r(v)$ the r-local q-type of v, i.e. the set of all r-local formulas $\psi(x)$ of quantifier rank q such that $G \models \psi(v)$.

▶ **Corollary 3.4.** *For every formula $\psi(x,y)$ there exist numbers r and q such that for every graph G the following holds: If u and v are two vertices of G such that the distance between them is more than $2r$, then whether $G \models \psi(u,v)$ depends only on $tp_q^r(u)$ and $tp_q^r(v)$.*

3.5 Graph classes

We assume familiarity with the notions of treewidth and of clique-width. We will need the following results about the latter concept.

▶ **Theorem 3.5** ([2]). *Let $\mathcal{C}$ be a class of graphs which is interpretable in a graph class of bounded treewidth. Then $\mathcal{C}$ is of bounded clique-width.*

▶ **Theorem 3.6** ([2]). *The FO model checking problem is solvable in fpt runtime on classes of graphs of bounded clique-width.*

We remark that Theorem 3.6 assumes that a clique-width decomposition of the input graph G is provided together with G and it is not known how to efficiently compute an optimal clique-width decomposition. However, one can approximate clique-width using the notion of rankwidth [24], and rankwidth decompositions can be efficiently computed [20].

Nowhere dense graph classes were introduced by Nešetřil and Ossona de Mendez [23]. Instead of working with the original definition, we will be working with an equivalent notion of *uniform quasi-wideness*. Informally, the definition says that we can obtain a large r-scattered set in any sufficiently large set $A \subseteq V(G)$ by removing a few vertices from G.

▶ **Definition 3.7** (Uniform quasi-wideness [4, 5]). *A class $\mathcal{C}$ of graphs is uniformly quasi-wide if for each $r \in \mathbb{N}$ there is a function $N : \mathbb{N} \to \mathbb{N}$ and a constant $s \in \mathbb{N}$ such that for every $k \in \mathbb{N}$, graph $G \in \mathcal{C}$ and subset A of $V(G)$ with $|A| \ge N(k)$, there is a set S of size $|S| \le s$ such that in $A \setminus S$ there are at least k vertices with pairwise distance more than r in $G \setminus S$.*

▶ **Theorem 3.8** ([22]). *A class $\mathcal{C}$ of graphs is uniformly quasi-wide if and only if $\mathcal{C}$ is nowhere dense.*

3.6 Parameterized complexity

We refer to [3] for an indepth introduction to parameterized complexity and only briefly recall the concepts from parameterized complexity needed below. A *parameterized problem* P is essentially a classical problem but in addition to the normal input instance w we are given an integer k, the so-called *parameter*. The problem P is called *fixed-parameter tractable*, or in the complexity class FPT, if there is a computable function f and a constant c such that the problem can be solved by an algorithm whose running time on input (w, k) is bounded by $f(k) \cdot |w|^c$. The class FPT can be seen as the parameterized equivalent to the classical complexity class P as abstraction of efficiently solvable problems.

4 Differential model checking

In this section, we show that in order to efficiently decide whether $G \models \varphi$, it is enough to efficiently solve the following problem: Given a labelled graph G, two of its vertices u and v, and a number q, decide whether there exists a formula $\psi(x)$ with quantifier rank q such that $G \models \psi(u)$ and $G \not\models \psi(u)$, i.e. whether $u \equiv_q^1 v$. Moreover, it is enough to compute a relation $\sim$ (not necessarily an equivalence) such that the transitive closure of $\sim$ is a refinement of $\equiv_q^1$ with the number of classes bounded in terms of q.

Due to the space restrictions, we only sketch the idea behind the main result of this section and refer to the full version for details.

First note that, if we can decide for two vertices u, v of G whether $u \equiv_q^1 v$ efficiently, then we can compute the whole relation $\equiv_q^1$ on $V(G)$ with quadratic overhead and pick a representative for each class of $\equiv_q^1$. Let $\varphi = Q_1x_1Q_2x_2\ldots, Q_qx_q\psi(x_1, x_2 \ldots, x_q)$ be a sentence in prenex normal form and let $v_1, \ldots, v_m$ be representatives of all classes of $\equiv_{q-1}^1$ on $V(G)$. Then the following is true in case $Q_1 = \exists$ we have: $G \models \varphi$ if and only if among $v_1, \ldots, v_m$ there is a v such that $G \models Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$. In case $Q_1 = \forall$ the following holds: $G \models \varphi$ if and only if for every v among $v_1, \ldots, v_m$ we have $G \models Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$. Thus, instead of going through every $v \in V(G)$ and evaluating $Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$ on G (the naive evaluation algorithm), we only need to evaluate $Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$ on every vertex v from $v_1, \ldots, v_m$, and m is bounded in terms of q. To evaluate $Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$ for any fixed v from $v_1, \ldots, v_m$, we proceed as follows: Mark v in G by label l_1 and all its neighbours by label l_1' to obtain graph G'. A simple argument shows that we can transform $\psi(x_1, x_2, \ldots, x_q)$ into $\psi'(x_2, \ldots, x_q)$ such that it holds that $G \models Q_2x_2\ldots, Q_qx_q\psi(v, x_2 \ldots, x_q)$ if and only if $G' \models Q_2x_2\ldots, Q_qx_q\psi'(x_2 \ldots, x_q)$. Thus, we are again left with a problem of evaluating a sentence $\varphi' = Q_2x_2\ldots, Q_qx_q\psi'(x_2 \ldots, x_q)$ on G', but this time the sentence has one less quantifier and the graph is labelled. If we can find representatives of all classes of $\equiv_{q-2}'$, where $\equiv_{q-2}'$ is taken over the original vocabulary extended by the two labels l_1 and l_1', we can continue in this fashion until we eliminate all quantifiers and evaluate the original sentence φ on G.

It is easily seen that the above approach also works if we can just compute a subrelation $\sim_q$ of $\equiv_q^1$, provided that the graph of $\sim_q$ has bounded number of connected components (in terms of q). Finally, if the graph of $\sim_q$ has bounded independence number, one can greedily compute the representatives of classes in the transitive closure of $\sim_q$ with linear overhead instead of quadratic one. This leads to the following corollary.

▶ **Corollary 4.1.** *Let $\mathcal{C}$ be a class of graphs with a function p such that for every t, every $G \in \mathcal{C}$ with at most t labels and every q there is a symmetric and reflexive relation $\sim_{q,t}$ on $V(G)$ such that*

1. *Every class of $\equiv_q^{1,G}$ (over the vocabulary extended with t labels) is a union of connected components of the graph of $\sim_{q,t}$ (in other words the transitive closure of $\sim_{q,t}$ is a refinement of $\equiv_q^{1,G}$),*
2. *The maximum size of any independent set in the graph of $\sim_{q,t}$ is bounded by $p(q,t)$, and*
3. *We can decide whether $u \sim_{q,t} v$ in time $|V(G)|^c \cdot h(q,t)$ for some function h.*

Then one can perform model checking on $\mathcal{C}$ for any sentence φ in prenex normal form, with quantifier rank q, in time $|V(G)|^{c+1} \cdot g(q)$ for some function g.

In Section 6, we show that for graph classes interpretable in nowhere dense graph classes it makes sense to consider as $\sim_q$ the relation "Duplicator wins the $l(q)$-round differential game between vertices u and v of G", for some function l, as this relation satisfies items 1 and 2 above. The situations when it also satisfies item 3 are discussed in sections 7 and 8.

5 Semi-differential EF game

Based on the results from Section 4, to perform model checking efficiently it is enough to be able to determine whether two vertices u and v of a graph G are of the same q-type. To this end we introduce *semi-differential EF games* (this section) and *differential EF games* (next section). The main differences compared to the standard EF game are that the (semi-)differential game is played only on one graph and the moves of the Spoiler (and also the Duplicator in the differential game) are restricted.

For two vertices u, v of a graph G we denote by $D(u,v)$ the symmetric difference of their neighbourhoods, which we will call their *differential neighbourhood*, i.e.

$$D(u,v) := N(u) \Delta N(v).$$

If the graph G in which we want to take the differential neighbourhood is not clear from the context, we will add the relevant graph as an index, as in $D^G(u,v)$.

▶ **Definition 5.1.** *The semi-differential EF game $\mathcal{G}_m^{SD}(G, a_1, \ldots, a_k, b_1, \ldots, b_k)$ with $m \in \mathbb{N}$ rounds is defined in the same way as the standard EF game with the following differences:*

1. *The game is played only on one graph G and $a_1, \ldots, a_k, b_1, \ldots, b_k$ all stem from $V(G)$.*[1]
2. *The starting position is $((a_1, \ldots, a_k), (b_1, \ldots, b_k))$.*
3. *In the j-th round the Spoiler is only allowed to make a move on a vertex $v \in D(a_i, b_i)$ for some $i < j + k$. The Spoiler decides whether this move defines a_j or b_j, i.e. whether the position after his move is $((a_1, \ldots, a_k, v), (b_1, \ldots, b_k))$ or $((a_1, \ldots, a_k), (b_1, \ldots, b_k, v))$. In case no such v exists, the Duplicator wins.*
4. *Duplicator's moves are unrestricted and her reply becomes b_j, if Spoiler decided that the vertex he chose becomes a_j, or a_j, otherwise.*

If $\mathcal{G}_m^{SD}(G, a_1, \ldots, a_k, b_1, \ldots, b_k)$ is won by the Duplicator, we write $a_1, \ldots, a_k \cong_m^{k,SD} b_1, \ldots, b_k$. We apply the same notational conventions to $\cong_m^{k,SD}$ as we did to $\cong_m^k$. If a vertex is played to append the tuple $(a_1, \ldots, a_k)$, we call it an *a-move*, otherwise we call it a *b-move*.

[1] One can also think of the game being played on two copies G_1 and G_2 of graph G, with $a_1, \ldots, a_k \in V(G_1)$ and $b_1, \ldots, b_k \in V(G_2)$. However, we need to be able to refer to $D(a_i, b_i)$, and this is more convenient if both a_i and b_i are in the same graph.

We will from now on refer to the semi-differential EF game as the *semi-differential game*. Let $\cong^{SD}_{m,G}$ denote the relation "Duplicator wins the m round differential game between u and v on the graph G". As usual, we will drop the index G if the graph is clear from the context.

We note that, due to the distinction between a-moves and b-moves, it is also possible that the subgraphs induced by these tuples are not connected. Consider for example a semi-differential game on P_4, with $V(P_4) = [4]$ and $E(P_4) = \{\{i, i+1\} \mid i \in [3]\}$, starting on the position $((a_1 = 1), (b_1 = 4))$. The Spoiler can now play $a_2 = 3$, since $D(a_1, b_1) = \{2, 3\}$, and the graph $P_4[\{a_1, a_2\}]$ is not connected.

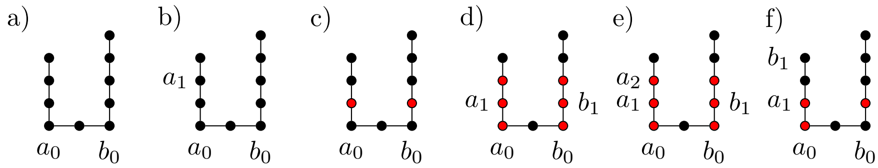

Figure 1 Pictures a) and b): A regular EF game on one graph. The starting position is depicted in a), and the position after Spoiler's first move in b). No matter where the Duplicator replies, she will lose in at most two more moves. Pictures c) - f): The semi-differential EF game. The starting position is depicted in c); all vertices into which the Spoiler can move are marked in red. In this case the Spoiler cannot play the same first move as in the regular EF game. However, he can still play this vertex in two moves (example play in d) and e)) and win from there. The Duplicator can prevent this by playing b_1 as in f) (her moves are unrestricted.), but this loses immediately, since $a_0 a_1 \in E$ but $b_0 b_1 \notin E$.

Semi-differential games have a direct relation to regular EF games. The rest of the section is devoted to proving that, at the cost of playing more moves, we can play a semi-differential game instead of a regular EF-game to distinguish two vertices. The proof idea is also partially illustrated in Figure 1.

▶ **Lemma 5.2.** *For every $m \in \mathbb{N}$ there exists $l = l(m) \in \mathbb{N}$ such that for every graph G it holds that if $\bar{a} \ncong_m \bar{b}$, then $\bar{a} \ncong^{SD}_{l(m)} \bar{b}$.*

Proof. We set $l(0) := 0$ and $l(i+1) := 2l(i) + 1$ and prove the claim by induction on m. For $m = 0$ there is nothing to prove. For the induction step, we assume that the claim holds for m and prove it for $m+1$. Let $\bar{a} = (a_1, \dots, a_k)$ and $\bar{b} = (b_1, \dots, b_k)$ be the starting position. Since by our assumptions the Spoiler has a winning strategy, there exists $v \in V(G)$ such that for every $u \in V(G)$ the Spoiler has a winning strategy in the m-round EF game from position $((\bar{a}, v), (\bar{b}, u))$. In particular, there exists a winning strategy for the Spoiler if $v = u$. By our induction hypothesis, the Spoiler wins the $l(m)$-round semi-differential game starting from $((\bar{a}, v), (\bar{b}, v))$. We fix the Spoiler's winning strategy S for this semi-differential game and apply it to the position $(\bar{a}, \bar{b})$. This is possible because $D(v, v) = \emptyset$. Let $((\bar{a}, a_{k+1}, \dots, a_{k+l(m)}), (\bar{b}, b_{k+1}, \dots, b_{k+l(m)}))$ be the position after $l(m)$ rounds. If the subgraphs of G induced by $(\bar{a}, a_{k+1}, \dots, a_{k+l(m)})$ and $(\bar{a}, b_{k+1}, \dots, b_{k+l(m)})$ are not isomorphic, then the Spoiler has already won. If they are isomorphic, then it has to hold that $v \in D(a_i, b_i)$, for some $i \in \{k+1, \dots, k+l(m)\}$, as otherwise the Duplicator's moves would beat the Spoiler's strategy S in the m-round semi-differential game starting from the position $((\bar{a}, v), (\bar{b}, v))$.

Since $v \in D(a_i, b_i)$, for some $i \in \{k+1, \dots, k+l(m)\}$, the Spoiler can play v in the next round. Let u be the Duplicator's reply. By our assumptions, the Spoiler wins the m-round EF game from the position $((\bar{a}, v), (\bar{b}, u))$. Therefore, according to the induction hypothesis, there exists a winning strategy for the Spoiler in the semi-differential game with $l(m)$ rounds from the position $((\bar{a}, v), (\bar{b}, u))$; let S' be this strategy. Since we only restrict the Spoiler's moves in the semi-differential game, applying S' to the position $((\bar{a}, a_{k+1}, \dots, a_{k+l(m)}, v), (\bar{b}, b_{k+1}, \dots, b_{k+l(m)}, u))$ will not change the outcome and thus the Spoiler wins. ◀

Even though we are not able to establish that the relation $\cong^{SD}_{m,G}$ is an equivalence, we get the following.

▶ **Lemma 5.3.** *For every $m \in \mathbb{N}$ and every graph G, the number of connected components of the graph of the relation $\cong^{SD}_{l(m)}$ is bounded by a function of m, and each equivalence class of $\cong_m$ is a union of connected components of the graph of $\cong^{SD}_{l(m)}$.*

6 Differential game

In the semi-differential game we restrict Spoilers moves to $D(a_i, b_i)$ but the Duplicator's moves are unrestricted. For the application we have in mind, we will restrict Duplicators moves to $D(a_i, b_i)$ as well, using the same i picked by the Spoiler, and call the resulting game the *differential game*. For every graph G, we define the relation $\cong^D_m$ on $V(G)$ by setting $u \cong^D_m v$ if and only if Duplicator wins the m-round differential game starting from u and v. We extend this notation to tuples of vertices in the same way as we did with the previous two game definitions.

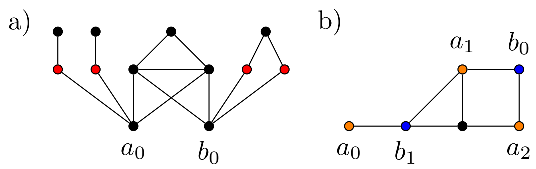

Figure 2 Example positions/plays of the differential game. Notice that in a) from the starting position no vertex of the upper middle triangle can be played by either player. Meanwhile in b), we started from a_0 and b_0, then played a_1, forcing the Duplicator's choice of b_1. Subsequently, the choice of a_2 produces an independent set of size three, which the Duplicator cannot replicate. It can therefore be useful to play a sequence of disconnected vertices even in the differential game.

We summarize the basic properties of differential games used in this section in the following lemma. We refer to the full version for the proof.

▶ **Lemma 6.1** (Properties of differential games). *Let l be the function from Lemma 5.2.*

1. *Every class of $\cong_m$ is a union of connected components of the graph of $\cong^D_{l(m)}$.*
2. *Let u and v be two vertices of a graph G such that $u \cong_m v$ and the distance between u and v is more than $2m$. Then the Duplicator wins the m round differential game between u and v.*
3. *For every $m \in \mathbb{N}$ and every graph G, if $\bar{a} \ncong_m \bar{b}$, then $\bar{a} \ncong^D_{l(m)} \bar{b}$.*
4. *Let G be a graph and let $\bar{a} := a_1, \dots, a_k$, $\bar{b} := b_1, \dots, b_k$ and w be vertices of G such that $\bar{a}w \ncong_m \bar{b}w$. Then the Spoiler has a strategy in $\mathcal{G}^D_{l(m)+1}(G, \bar{a}, \bar{b})$ such that he can play w at some point or he wins $\mathcal{G}^D_{l(m)+1}(G, \bar{a}, \bar{b})$.*
5. *Let $\psi(x, y)$ be an interpretation formula of quantifier rank q and let G and H be graphs such that $H = I_\psi(G)$. Let a and b be two vertices of G such that $a \cong^D_{(m+1)(l(q)+1)} b$ in G. Then $a \cong^D_m b$ in H.*

We now show that for any first order interpretation $\psi(x,y)$ and any nowhere dense (uniformly quasi-wide) class $\mathcal{C}$ of graphs the number of components of $\cong_m^D$ in any $G \in I_\psi(\mathcal{C})$ is bounded.

▶ **Theorem 6.2.** *Let $\mathcal{C}$ be a uniformly quasi-wide class of labelled graphs and let $\psi(x,y)$ be a first order interpretation formula. Then for each m there exists p such that for every $H \in I_\psi(\mathcal{C})$ the maximum size of any independent set in the graph of $\cong_m^{D,H}$ is at most p.*

Proof. Suppose that there exists an m such that for every p there is a graph $H \in I_\psi(\mathcal{C})$ for which there is an independent set of size more than p in the graph of $\cong_m^{D,H}$.

In what follows we assume that the graph H we work with is as large as necessary and the maximum size of an independent set in the graph of $\cong_m^D$ is as large as we need in our argumentation. We set A to be an independent set of maximum size in the graph of $\cong_m^D$. We therefore have $|A| > p$, and since we can choose p to be arbitrarily large, we can ensure that $|A|$ is as large as we want.

Let q be the quantifier rank of ψ and set $d \coloneqq (m+1)(l(q)+1)$, where l is the function from Lemma 5.2. Since $H \in I_\psi(\mathcal{C})$, there exists at least one $G \in \mathcal{C}$ such that $I_\psi(G) = H$. Because $\mathcal{C}$ is uniformly quasi-wide, we know (by applying the definition to $r = 2d+1$) that there exists a constant s and a function M such that for any number k and any set A of size at least $M(k)$, it is possible to remove s vertices from G such that there are k vertices $v_1, \ldots, v_k$ at pairwise distance at least r in $G \setminus S$. We create the graph G' from $G \setminus S$ by putting the vertices from S back (but without any edges) and labelling them each with a different colour from $[s]$. Additionally, we label the neighbourhood in G of each vertex with colour i with the label l_i. In G' the vertices $v_1, \ldots, v_k$ remain pairwise at distance more than $2m$. We can recover G from G' with a quantifier-free interpretation $\delta(x,y)$. By concatenating δ and ψ, we obtain an interpretation formula $\psi'(x,y)$, with quantifier rank q, such that $H = I_{\psi'}(G')$.

We choose A to be large enough so that among $v_1, \ldots, v_k$ there exists a pair of distinct vertices v_a and v_b with $v_a \cong_d^{G'} v_b$. Note that k only has to be larger than the number of classes of relation $\cong_r^1$ with respect to the vocabulary of $\mathcal{C}$ enriched by $2s$ labels, and so k does not depend on the particular graph G. Since v_a and v_b lie at distance r in G' and $r > 2d$, we can use part 2 of Lemma 6.1 to conclude that $v_a \cong_d^{D,G'} v_b$ is true as well. We can now use part 5 of Lemma 6.1 to conclude that $v_a \cong_m^{D,H} v_b$, which contradicts our assumptions because v_a and v_b come from an independent set of the graph of $\cong_m^{D,H}$. ◀

Combining part 1 of Lemma 6.1 and Theorem 6.2 with Corollary 4.1 we get the following theorem.

▶ **Theorem 6.3.** *Let $\mathcal{C}$ be a class of labelled graphs interpretable in a nowhere dense class of graphs such that we can decide the winner of the m-round differential game in fpt runtime with respect to the parameter m. Then the FO model checking problem is solvable in fpt runtime on $\mathcal{C}$.*

Proof. Let $\mathcal{C}$ be a class of graphs with labels from the set $\{1, \ldots, t\}$ and properties assumed in the statement of the theorem. From part 1 of Lemma 6.1 it follows that for every $G \in \mathcal{C}$ it holds that the closure of $\cong_{l(q)}^{D,G}$ refines $\cong_q^{1,G}$ and by Theorem 3.2 the relation $\cong_q^1$ is the same as $\equiv_q^1$. It follows that the closure of $\cong_{l(q)}^{D,G}$ refines $\equiv_q^{1,G}$. By Theorem 6.2 we know that the maximum size of and independent set in the graph of $\equiv_{l(q)}^{D,G}$ is bounded in terms of $l(q)$. It follows that we can use $\cong_{l(q)}^{D,G}$ as $\sim_{q,t}$ in Corollary 4.1. ◀

7 Applications

Theorem 6.3 from the previous section immediately implies the well-known result that FO model checking can be solved in fpt runtime on classes of graphs of bounded degree. In fact, one can easily use it to establish the (already known) existence of FO model checking algorithms for locally simple graph classes, such as graph classes of locally bounded treewidth – in this case the whole m-round game between any two vertices u, v will take place in $N_m(u) \cup N_m(v)$, and $G[N_m(u) \cup N_m(v)]$ has bounded treewidth, so one can decide the winner there efficiently. The more complicated case of graph classes interpretable in graphs with locally bounded treewidth is analysed in Section 8.

Returning to the case of graphs of bounded degree, the following theorem was proven in [16].

▶ **Theorem 7.1.** *Let $\mathcal{C}$ be a class of graphs of maximum degree d, ψ an interpretation formula and $\mathcal{D} = I_\psi(\mathcal{C})$. Then we can perform FO model checking on $\mathcal{D}$ in fpt runtime.*

The idea behind the proof of Theorem 7.1 in [16] was to compute an approximate "reversal" of the interpretation: Given $H \in \mathcal{D}$, one computes in fpt runtime a graph G of degree d' and formula $\psi'(x, y)$ such that $H = I_{\psi'}(G)$ and such that both d' and ψ' depend only on d and ψ. From this one can then establish Theorem 7.1 by standard considerations.

One can consider a more general version of the approximate interpretation reversal mentioned above. Let $\mathcal{D}$ be a class of graphs interpretable in a nowhere dense class of graphs $\mathcal{C}$ using formula $\psi(x, y)$. The task is to find a nowhere dense class $\mathcal{C}'$ of graphs and a polynomial time algorithm which, given $H \in \mathcal{D}$ as input, computes in a graph $G \in \mathcal{C}'$ and formula $\psi'(x, y)$ such that $H = I_{\psi'}(G)$ and such that ψ' depends only on ψ. A general solution to this problem would imply that there is an fpt FO model checking algorithm for every class of graphs interpretable in a nowhere dense graph class. Unfortunately, the problem of efficiently computing interpretation reversals seems to be currently out of reach, and only small progress has been made so far.

Our techniques suggest that in some cases it may be possible to avoid computing approximate interpretation reversals and efficiently model check formulas directly on the class of graphs interpreted in a class of sparse graphs. To illustrate this, we now reprove Theorem 7.1 in this fashion. We will need the following definition.

▶ **Definition 7.2.** *A pair u, v of vertices of a graph G is* (m, k)-good *if the Duplicator has a winning strategy in the m-round differential game starting from $a_0 := u, b_0 := v$ such that, for any pair a_i, b_i played in the course of the game, it holds that $|D(a_i, b_i)| \leq k$*

Intuitively, the notion of (m, k)-goodness captures the situation when the Duplicator can win the m-round Differential game between u and v in such a way that the "arena" of admissible moves never gets too big. This is useful because one can easily check in fpt runtime whether two vertices are (m, k)-good by playing the differential game by brute-force and whenever for some a_i, b_i it holds that $|D(a_i, b_i)| > k$, declare the current branch as lost for Duplicator.

▶ **Proposition 7.3.** *Given two vertices u, v of a graph G, one can check in time $f(m, k) \cdot |V(G)|^2$ whether they are (m, k)-good.*

With the notion of (m, k)-goodness one can prove Theorem 7.1 easily by showing that the relation $\simeq_{m,k}$ defined by "$u \simeq_{m,k} v$ if and only if u and v are (m, k)-good" can be taken as $\sim_{q,t}$ in Corollary 4.1 for suitable values of m and k (depending on q, t, ψ, d). We refer to the full version for details.

8 Differentially simple graph classes

Based on the results from the previous sections, in order to evaluate FO sentences in prenex normal form on a graph class $\mathcal{C}$ interpretable in a nowhere dense graph class, it is enough to be able to determine the winner of the differential game on pairs of vertices of a graph from $\mathcal{C}$. This can, however, be too difficult – for example consider the case in which $\mathcal{C}$ is an interpretation of planar graphs. In this case for many pairs of vertices u, v of G from $\mathcal{C}$ it can happen that $D(u, v)$ contains most of, or even the entire, graph. We can sidestep this problem by giving such vertices u and v different labels. This essentially means that whenever the Duplicator would play u as a reply to v (or vice versa), she would already have lost from that point on, and so $D(u, v)$ would be irrelevant. Extending these ideas to more than one round leads to the following definitions.

▶ **Definition 8.1** (Differential neighbourhoods). *Let G be a coloured graph and let $c(a)$ denote the colour of a vertex a of G.*

- *The differential 1-neighbourhood $DN_1(u, v)$ of vertices u, v with $c(u) = c(v)$ is the set $D(u, v)$.*
- *For $r \in \mathbb{N}$ with $r > 1$,*

$$DN_r(u, v) := \bigcup_{\substack{a,b \in DN_{r-1}(u,v) \\ c(a)=c(b)}} D(a, b) \cup DN_{r-1}(u, v).$$

- *For $r \in \mathbb{N}$, the* closed *differential r-neighbourhood is defined as $DN_r[u, v] := DN_r(u, v) \cup \{u, v\}$.*

▶ **Definition 8.2.** *We say that class $\mathcal{C}$ is* differentially simple *if for every r there exists $m_r \in \mathbb{N}$ and a graph class $\mathcal{D}_r$ with an efficient FO model checking algorithm such that it is possible to colour every $G \in \mathcal{C}$ with m_r colours such that for every pair u, v of vertices of the same colour it holds that $G[DN_r[u, v]]$ is a graph from $\mathcal{D}_r$.*

Note that if $\mathcal{C}$ is interpretable in a nowhere dense class of graphs, then adding at most m labels to each graph from $\mathcal{C}$ does not change the fact that the maximum size of an independent set in the graph of $\cong_r^D$ for each G from $\mathcal{C}$ is bounded, because Theorem 6.2 works with labelled graphs. We can thus focus on determining the winner of the differential game on $G[DN_r[u, v]]$, which is from $\mathcal{D}_r$, instead of on G which is from $\mathcal{C}$. In case $\mathcal{D}_r$ is a class of graphs with efficient model checking algorithm, we can use this to determine the winner of the game. We will use the FO formula $\xi_r(x, y)$, which expresses that Duplicator wins the r-round differential game between x and y (it is easy to construct such formula for each r), and evaluate it on $G[DN_r[u, v]]$ using the model checking algorithm for $\mathcal{D}_r$. The following theorem summarises this.

▶ **Theorem 8.3.** *Let $\mathcal{C}$ be a differentially simple class of graphs such that*

- *$\mathcal{C}$ is interpretable in a nowhere dense class of graphs, and*
- *There exists an fpt algorithm (with respect to the parameter r) which computes the colouring from Definition 8.2 for every $G \in \mathcal{C}$.*

Then the FO model checking problem is in FPT on $\mathcal{C}$.

To show that differentially simple graph classes can be useful, we prove that classes of graphs interpretable in graph classes with locally bounded treewidth are differentially simple, where each graph class $\mathcal{D}_r$ is a class of graphs of bounded clique-width. We note that in this case there is one m-colouring which works for every value of r and which satisfies the requirements of Definition 8.2.

▶ **Lemma 8.4.** *Let $\mathcal{C}$ be a class of graphs which is interpretable in a class of graphs of locally bounded treewidth. Then $\mathcal{C}$ is differentially simple.*

Proof. Let $\mathcal{E}$ be a class of graphs of locally bounded treewidth and $\psi(x, y)$ an interpretation formula such that $\mathcal{C} = I_\psi(\mathcal{E})$. From Gaifman's theorem applied to $\psi(x, y)$ it follows (by Corollary 3.4) that there exist d and q such that the following holds for any $G \in \mathcal{E}$ and $H \in \mathcal{C}$ such that $H = I_\psi(G)$. If u, v are two vertices of the same d-local q-type in $G \in \mathcal{E}$ and the vertex w is at distance more than $2d$ from both u and v in G, then $G \models \psi(u, w)$ iff $G \models \psi(v, w)$, which in turn means that $uw \in E(H)$ iff $vw \in E(H)$. It follows that if u, v are two vertices of H such that in G these vertices have the same d-local q-types, then every vertex in $D^H(u, v)$ has to be in the $2d$-neighbourhood of u or v in G.

We define the colouring of any $H \in \mathcal{C}$ as follows. Let $G \in \mathcal{E}$ be such that $H = I_\psi(G)$. We colour every vertex v of H by its d-local q-type in G. By the above considerations for any two vertices $u, v \in V(H)$ of the same colour it has to hold that every vertex in $D^H(u, v)$ has to come from $N^G_{2d}(u) \cup N^G_{2d}(v)$. If we consider any two vertices $u', v' \in D^H(u, v)$ of the same colour, the same argumentation applies – every vertex w in $D^H(u', v')$ has to come from $N^G_{2d}(u') \cup N^G_{2d}(v')$ and thus has to be at distance at most $2d$ from u or v in G, which means $w \in N^G_{2d}(u) \cup N^G_{2d}(v)$. It follows by an easy inductive argument that $DN^H_r(u, v)$ in H is a subset of $N^G_{2dr}(u) \cup N^G_{2dr}(v)$ for any positive integer r. Since $\mathcal{E}$ is a class of graphs of locally bounded treewidth, the subgraph of G induced by $N^G_{(r+1)2d}[u] \cup N^G_{(r+1)2d}[v]$ has treewidth bounded in terms of $(r + 1)2d$ and an easy argument shows that $H[DN^H_r[u, v]]$ is an induced subgraph of $I_\psi(G[N^G_{(r+1)2d}[u] \cup N^G_{(r+1)2d}[v]])$ which has bounded clique-width. ◀

Lemma 8.4 implies that if we are able to efficiently compute the colourings from Definition 8.2, then we obtain an efficient FO model checking algorithm for classes of graphs interpretable in graph classes of locally bounded treewidth by means of Theorem 8.3. However, the existence of such a colouring algorithm is unknown.

9 Discussion and open problems

We have introduced the notions of differential games and differential locality, which can lead to efficient model checking algorithms and which seem to be more "interpretation friendly" than Gaifman's theorem. We believe that the ideas outlined in this paper can lead to improved understanding of the structure of graphs interpretable in sparse graphs, and perhaps also lead to some insights into stable graphs (if Theorem 6.2 gets strengthened to stable graph classes).

9.1 Complement-simple graph classes

Regarding our application to the model checking problem for graph classes interpretable in classes of graphs with locally bounded treewidth, it has to be noted that there exists a simpler approach based on colourings and on Gaifman's theorem, which avoids differential techniques altogether.

▶ **Definition 9.1.** *We say that a class $\mathcal{C}$ of graphs is* complement-simple *if for every r there exists m_r and graph class $\mathcal{D}_r$ with efficient FO model checking algorithm such that every $G \in \mathcal{C}$ has a m_r-colouring such that complementing edges between some pairs of colours results in a graph G' in which for every $v \in V(G')$ it holds that $N^G_r[v] \in \mathcal{D}_r$.*

If $\mathcal{C}$ is a complement-simple graph class such that we can compute colourings from Definition 9.1 efficiently, then we can perform FO model checking on graphs from $\mathcal{C}$ efficiently. To do this, note that we can interpret G in G' and that we can solve the model checking problem on G' efficiently by using Gaifman's theorem.

Coming back to graph classes interpretable in graph classes with locally bounded treewidth, using the colouring used in the proof of Lemma 8.4 one can show that every such graph class $\mathcal{C}$ is complement-simple and again one can use the same colouring for all values of r.

▶ **Proposition 9.2.** *Let $\mathcal{C}$ be a class of graphs which is interpretable in a class of graphs of locally bounded treewidth. Then $\mathcal{C}$ is complement-simple.*

Proof sketch. We will use the same colouring as was used in the proof of Lemma 8.4 and show that complementing the edges in G between some pairs of colours in this colouring leads to a graph G' with locally bounded clique-width. Let $H \in \mathcal{C}$ and let G be such that $H = I_\psi(G)$ and colour each vertex of H by its d-local q-type, where d and q come from Gaifman's theorem applied to $\psi(x, y)$. We say that an edge uv in $E(H)$ is *long* if $dist_G(u, v) > 2d$. Let $t_1 := tp^d_q(u)$ and $t_2 := tp^d_q(v)$. By Corollary 3.4, if there is a long edge in H between any two vertices of types t_1 and t_2, then there exists an edge between all pairs of vertices of type t_1 and t_2 which are at distance more than $2d$ in G. In this case we say that types t_1 and t_2 *induce long edges*. By complementing the edges between any pair of types (colours) in H which induce long edges we remove all long edges in H and obtain graph H'. It is easily shown that H' is interpretable in G (equipped with colours) by an interpretation which acts only locally and thus H' has locally bounded clique-width. ◀

9.2 Open problems

We conclude with several open problems and possible directions for future research.

1. Is it true that for any stable class $\mathcal{C}$ of graphs and any q there exists p such that every independent set in the graph of the relation $\cong^D_q$ has size at most p?
2. Is it possible to extend the ideas from the proof of Theorem 7.1 from Section 7 to more general graph classes? In particular, is it possible to use an approach based on differential games to design efficient algorithms on interpretations of sparse graphs that do not rely on computing approximate reversals of interpretations?
3. Let $\mathcal{C}$ be a class of graphs interpretable in graph classes of locally bounded treewidth. Is there a polynomial time algorithm that, for every $G \in \mathcal{C}$, computes a colouring such that for every r and every $u, v \in V(G)$ the graph $G[DN_r[u, v]]$ has small clique-width (depending on r)? If not, is there an fpt algorithm which computes such a colouring for every r?
4. Is it possible to use an approach based on differential games to give simpler/different algorithms for the FO model checking problem on graph classes of bounded expansion or nowhere dense graph classes than the algorithms presented in [7] and [19]? If yes, is it possible to use it to extend these results to interpretations of nowhere dense graph classes?
5. More generally, if the answer to Question 1 is yes, is it possible to use our methods to attack the FO model checking problem on stable graph classes?
6. Is there a useful normal form for FO formulas (say, similar to Gaifman normal form) based on differential neighbourhoods and the formulas ξ_r?

References

1 É. Bonnet, E. J. Kim, S. Thomassé, and R. Watrigant. Twin-width i: tractable fo model checking. In *FOCS'20*, pages 601–612. IEEE Computer Society, 2020. doi:10.1109/FOCS46700.2020.00062.

2 B. Courcelle, J. A. Makowsky, and U. Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory Comput. Syst.*, 33(2):125–150, 2000. doi:10.1007/s002249910009.

3 Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. doi:10.1007/978-3-319-21275-3.

4 A. Dawar. Finite model theory on tame classes of structures. In *International Symposium on Mathematical Foundations of Computer Science*, pages 2–12. Springer, 2007. doi:10.1007/978-3-540-74456-6_2.

5 A. Dawar. Homomorphism preservation on quasi-wide classes. *Journal of Computer and System Sciences*, 76(5):324–332, 2010. doi:10.1016/j.jcss.2009.10.005.

6 A. Dawar, M. Grohe, and S. Kreutzer. Locally excluding a minor. In *LICS'07*, pages 270–279. IEEE Computer Society, 2007. doi:10.1109/LICS.2007.31.

7 Z. Dvorák, D. Kráľ, and R. Thomas. Testing first-order properties for subclasses of sparse graphs. *J. ACM*, 60(5):36:1–36:24, 2013. doi:10.1145/2499483.

8 H. Ebbinghaus and J. Flum. *Finite model theory*. Perspectives in Mathematical Logic. Springer, 1995.

9 A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Math*, 49(129-141):13, 1961.

10 K. Eickmeyer and K. Kawarabayashi. FO model checking on map graphs. In *FCT 2017*, volume 10472 of *Lecture Notes in Computer Science*, pages 204–216. Springer, 2017. doi:10.1007/978-3-662-55751-8_17.

11 R. Fraïssé. Sur une nouvelle classification des systemes de relations. *Comptes Rendus Hebdomadaires Des Seances De L' Academie Des Sciences*, 230(11):1022–1024, 1950.

12 R. Fraïssé. *Sur quelques classifications des systemes de relations*. PhD thesis, Université de Paris, 1955.

13 M. Frick and M. Grohe. Deciding first-order properties of locally tree-decomposable structures. *J. ACM*, 48(6):1184–1206, 2001. doi:10.1145/504794.504798.

14 H. Gaifman. On local and non-local properties. In *Proceedings of the Herbrand Symposium*, volume 107 of *Stud. Logic Found. Math.*, pages 105–135. Elsevier, 1982.

15 J. Gajarský, P. Hliněný, D. Lokshtanov, J. Obdržálek, S. Ordyniak, M. S. Ramanujan, and S. Saurabh. FO model checking on posets of bounded width. In *FOCS'15*, pages 963–974. IEEE Computer Society, 2015. doi:10.1109/FOCS.2015.63.

16 J. Gajarský, P. Hliněný, J. Obdržálek, D. Lokshtanov, and M. S. Ramanujan. A new perspective on FO model checking of dense graph classes. In *LICS '16*, pages 176–184. ACM, 2016. doi:10.1145/2933575.2935314.

17 J. Gajarský, S. Kreutzer, J. Nešetřil, P. Ossona de Mendez, M. Pilipczuk, S. Siebertz, and S. Torunczyk. First-order interpretations of bounded expansion classes. In *ICALP 2018*, volume 107 of *LIPIcs*, pages 126:1–126:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ICALP.2018.126.

18 R. Ganian, P. Hliněný, D. Kráľ, J. Obdržálek, J. Schwartz, and J. Teska. FO model checking of interval graphs. *Log. Methods Comput. Sci.*, 11(4:11):1–20, 2015.

19 M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. *J. ACM*, 64(3):17:1–17:32, 2017. doi:10.1145/3051095.

20 P. Hliněný and S. Oum. Finding branch-decompositions and rank-decompositions. *SIAM J. Comput.*, 38(3):1012–1032, 2008. doi:10.1137/070685920.

21 P. Hliněný, F. Pokrývka, and B. Roy. FO model checking on geometric graphs. *Comput. Geom.*, 78:1–19, 2019. doi:10.1016/j.comgeo.2018.10.001.

22 J. Nešetřil and P. Ossona de Mendez. First order properties on nowhere dense structures. *J. Symb. Log.*, 75(3):868–887, 2010. doi:10.2178/jsl/1278682204.

23 J. Nešetřil and P. Ossona de Mendez. On nowhere dense graphs. *Eur. J. Comb.*, 32(4):600–617, 2011. doi:10.1016/j.ejc.2011.01.006.

24 S. Oum and P. D. Seymour. Approximating clique-width and branch-width. *J. Comb. Theory, Ser. B*, 96(4):514–528, 2006. doi:10.1016/j.jctb.2005.10.006.

25 D. Seese. Linear time computable problems and first-order descriptions. *Math. Structures Comput. Sci.*, 6(6):505–526, 1996.

26 L. J. Stockmeyer. *The complexity of decision problems in automata theory and logic*. PhD thesis, Massachusetts Institute of Technology, 1974.

27 M. Y. Vardi. The complexity of relational query languages. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pages 137–146, 1982.