First-Order Model-Checking

Sebastian Siebertz

UNIVERSITY OF WARSAW

The Parameterized Complexity Summer School

Vienna, 03.09.2017

Model-checking

The model-checking problem

Let $\mathcal{L}$ be a logic and let $\mathcal{C}$ be a class of structures.

$\mathrm{MC}(\mathcal{L},\mathcal{C})$	
Input:	Structure $\mathfrak{A} \in \mathcal{C}$, formula $\varphi \in \mathcal{L}$
Problem:	Does $\mathfrak{A}$ satisfy φ, in symbols, $\mathfrak{A} \models \varphi$?

First-order logic FO

- Subgraph isomorphism for pattern graph H (on vertex set $\{1, \ldots, n\}$):

$$\exists x_1 \ldots \exists x_n \Big(\bigwedge_{ij \in E(H)} E(x_i, x_j) \wedge \bigwedge_{ij \notin E(H)} \neg E(x_i, x_j) \Big)$$

- Dominating set of size at most k:

$$\exists x_1 \ldots \exists x_k \forall y \Big(\bigvee_{1 \le i \le k} (y = x_i \vee E(y, x_i)) \Big)$$

Algorithmic meta-theorems

- Many computational problems can be described elegantly in logics.

Algorithmic meta-theorem

Every problem definable in a given logic $\mathcal{L}$ is tractable.

- Provide uniform explanation why problems are tractable.
- Establish general algorithmic techniques for solving them.
- Corresponding intractability results for logics exhibit natural boundaries beyond which these techniques fail.

Complexity of first-order model-checking

- $\Sigma_0 = \Pi_0$: quantifier-free first-order formulas.
- For $t \geq 0$, let Σ_{t+1} be the set of all formulas

$$\exists x_1 \dots \exists x_k \varphi, \quad \text{where } \varphi \in \Pi_t.$$

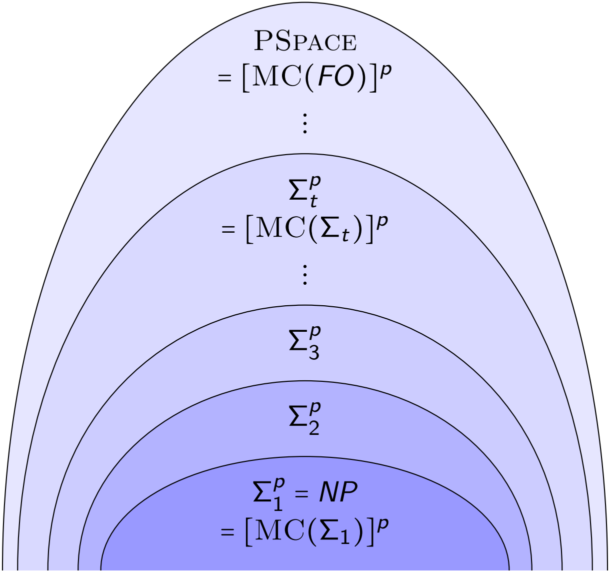
- For $t \geq 0$, let Π_{t+1} be the set of all formulas

$$\forall x_1 \dots \forall x_k \varphi, \quad \text{where } \varphi \in \Sigma_t.$$

Example

- $\exists x_1 \dots \exists x_n \big(\bigwedge_{ij \in E(H)} E(x_i, x_j) \wedge \bigwedge_{ij \notin E(H)} \neg E(x_i, x_j) \big) \in \Sigma_1$.
- $\exists x_1 \dots \exists x_k \forall y \big(\bigvee_{1 \leq i \leq k} (y = x_i \vee E(y, x_i)) \big) \in \Sigma_2$.

Complexity of first-order model-checking – the polynomial hierarchy



Parameterized model-checking

The parameterized model-checking problem

Let $\mathcal{L}$ be a logic and let $\mathcal{C}$ be a class of structures.

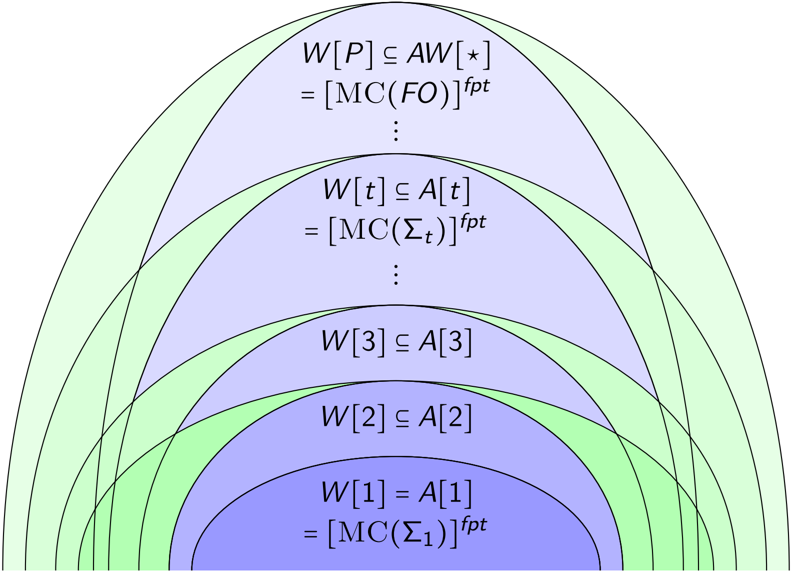
$\mathrm{MC}(\mathcal{L},\mathcal{C})$	
Input:	Structure $\mathfrak{A} \in \mathcal{C}$, formula $\varphi \in \mathcal{L}$
Parameter:	$\lvert\varphi\rvert$
Problem:	Does $\mathfrak{A}$ satisfy φ, in symbols, $\mathfrak{A} \models \varphi$?

- For $t \geq 1$ let $\Sigma_{t,1}$ be the set of all Σ_t formulas such that all quantifier blocks after the leading existential block have length ≤ 1.

Example

- $\exists x_1 \dots \exists x_n \big(\bigwedge_{ij \in E(H)} E(x_i, x_j) \wedge \bigwedge_{ij \notin E(H)} \neg E(x_i, x_j) \big) \in \Sigma_{1,1} = \Sigma_1$.
- $\exists x_1 \dots \exists x_k \forall y \big(\bigvee_{1 \leq i \leq k} (y = x_i \vee E(y, x_i)) \big) \in \Sigma_{2,1}$.

Parameterized complexity of first-order model-checking – *W*- and *A*-hierarchy



Parameterized complexity of first-order model-checking

Dominating Set is $W[2]$-hard.

Dominating Set

Input:	Graph G and $k \in \mathbb{N}$
Parameter:	k
Problem:	Do there exist k vertices which dominate G?

Clique-Dominating Set is $A[2]$-hard.

Clique-Dominating Set

Input:	Graph G and $k, \ell \in \mathbb{N}$
Parameter:	$k + \ell$
Problem:	Do there exist k vertices which dominate every clique of size ℓ?

Algorithmic meta-theorems

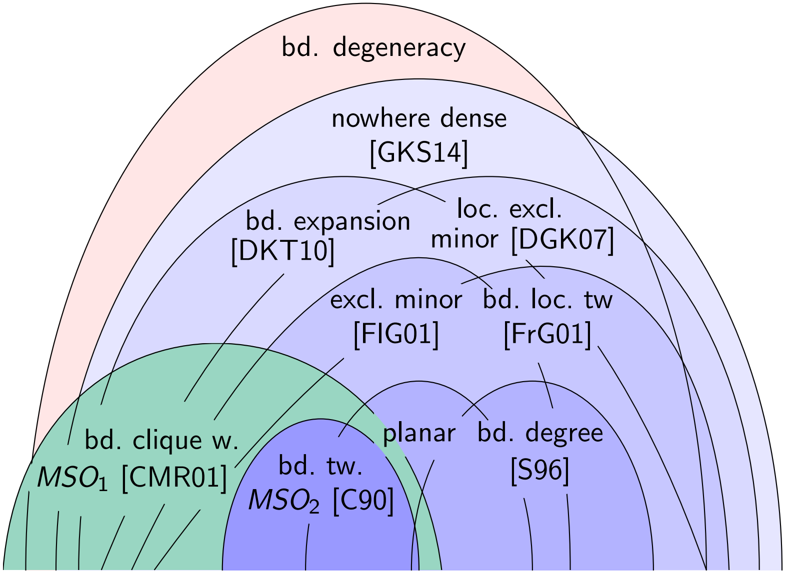
- Many computational problems can be described elegantly in logics.

Algorithmic meta-theorem

Every problem definable in a given logic $\mathcal{L}$ is tractable on every class of structures satisfying a certain property.

- Provide a uniform explanation why natural classes of problems are tractable on a certain class of structures (which may be sufficient for practical applications).
- Establish general algorithmic techniques for solving them.
- Corresponding intractability results for logics exhibit natural boundaries beyond which these techniques fail.

Sparse graph classes with fpt model-checking



Sparse graph classes with fpt model-checking

Theorem [Grohe, Kreutzer, S. 14]

If $\mathcal{C}$ is nowhere dense, then $\mathrm{MC}(FO)$ can be solved in time $f(|\varphi|) \cdot n^{1+\varepsilon}$ on every n-vertex graph $G \in \mathcal{C}$.

Theorem [Kreutzer 09], [Dvořák, Král', Thomas 11]

If $\mathcal{C}$ is somewhere dense and closed under taking subgraphs, then $\mathrm{MC}(FO)$ on $\mathcal{C}$ is $AW[\star]$-complete.

Corollary (assuming FPT $\neq AW[\star]$)

If $\mathcal{C}$ is closed under taking subgraphs, then $\mathrm{MC}(FO)$ on $\mathcal{C}$ is fpt if and only if $\mathcal{C}$ is nowhere dense.

Methods for sparse graphs

Theorem [Grohe, Kreutzer, S., 14]

If $\mathcal{C}$ is nowhere dense, then $\mathrm{MC}(FO)$ can be solved in time $f(|\varphi|) \cdot n^{1+\varepsilon}$ on every n-vertex graph $G \in \mathcal{C}$.

Methods

- Gaifman's theorem: $\varphi(x)$ equivalent to Boolean combination of local formulas and basic local formulas

$$\exists x_1 \ldots \exists x_k \bigwedge_{i \neq j} dist(x_i, x_j) > 2r \wedge \bigwedge_i \psi^{(r)}(x_i).$$

- Find r-neighbourhoods in which ψ is true (r only depends on φ).
- Solve distance-$2r$ independent set problem (k depends only on φ).
- Evaluate Boolean combination.

Lower bounds on somewhere dense graphs

Theorem

Let $\mathcal{C}$ be somewhere dense and closed under taking subgraphs. Then there exists $p \in \mathbb{N}$ such that for every graph H we have $H^p \in \mathcal{C}$, where H^p is the p-subdivision of H.

Proposition

Model checking on H^p is just as hard as on H.

Model-checking beyond sparse graphs

Corollary (assuming FPT $\neq AW[\star]$)

If $\mathcal{C}$ is closed under taking subgraphs, then $\mathrm{MC}(FO)$ on $\mathcal{C}$ is fpt if and only if $\mathcal{C}$ is nowhere dense.

Research program

Find the most general classes (which are not closed under taking subgraphs) which admit fpt model-checking.

- Efficient FO-model checking on specific dense graph classes.
  - Model-checking on certain interval graphs is fpt [Ganian et al. 13].
  - Model-checking on bounded width posets is fpt [Gajarský et al. 15].

→ There is no clear candidate for a *most general* dense class with tractable model-checking.

Model-checking beyond sparse graphs – Interpretations

- Assume $\mathcal{C}$ is a class with $\mathrm{MC}(FO, \mathcal{C}) \in \mathsf{FPT}$.
- For a graph G let $\bar{G}$ be its complement and let $\bar{\mathcal{C}} = \{\bar{G} : G \in \mathcal{C}\}$.

Proposition

Model-checking on $\bar{\mathcal{C}}$ is fpt.

Proof.

- Given $\bar{G} \in \bar{\mathcal{C}}$ and $\varphi \in FO$, let φ' be obtained from φ by replacing each atom $E(x, y)$ by $\neg E(x, y)$.
- Then $\bar{G} \models \varphi \iff G \models \varphi'$.
- G and φ' are efficiently computable from $\bar{G}$ and φ.

Model-checking beyond sparse graphs – Interpretations

- A *simple interpretation* $\mathcal{J}_\varphi$ is a formula $\varphi(x, y)$.
- For a graph $G = (V, E)$ define

$$\mathcal{J}_\varphi(G) = (V, \{uv : G \models \varphi(u, v) \vee \varphi(v, u)\}).$$

Example

On the previous slide we had $\varphi = \neg E(x, y)$ and $\mathcal{J}_\varphi(G) = \bar{G}$.

Interpretation Lemma

Replace in a formula ψ every occurrence of $E(x, y)$ by $(\varphi(x, y) \vee \varphi(y, x))$ to obtain $\mathcal{J}_\varphi(\psi)$. Then for every graph G

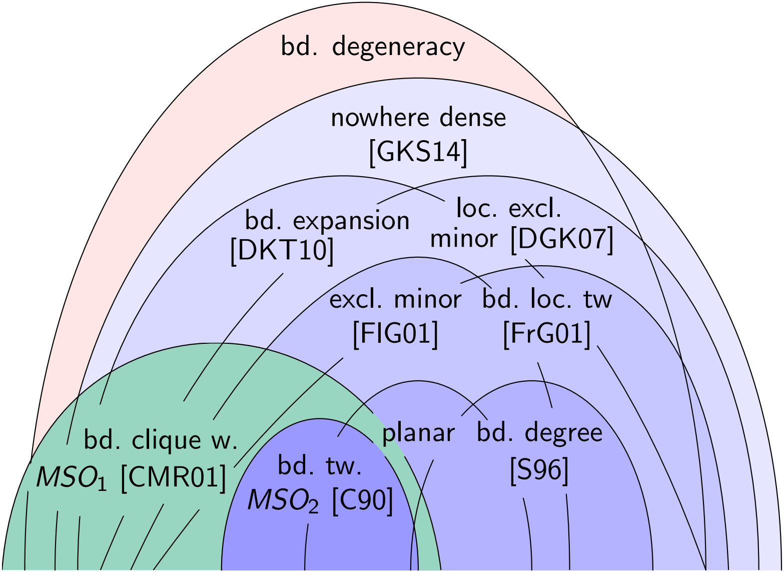
$$\mathcal{J}_\varphi(G) \models \psi \iff G \models \mathcal{J}_\varphi(\psi).$$

Model-checking beyond sparse graphs – Interpretations

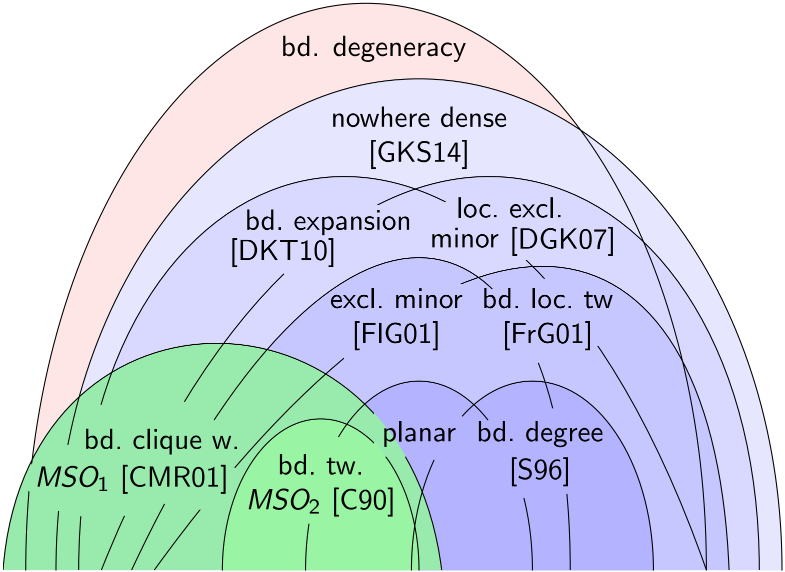
- Let $\mathcal{C}$ be a class with efficient model-checking.
- Let $\varphi(x,y)$ be a formula and $\mathcal{J}_\varphi(\mathcal{C}) = \{\mathcal{J}_\varphi(G) : G \in \mathcal{C}\}$.
  - On input $H = \mathcal{J}_\varphi(G) \in \mathcal{J}_\varphi(\mathcal{C})$ and ψ:
  - Compute $\mathcal{J}_\varphi(\psi)$.
  - We have $H \models \psi \iff G \models \mathcal{J}_\varphi(\psi)$.
  - $G \models \mathcal{J}_\varphi(\psi)$ can be decided efficiently.

Problem: We do not know how to compute G from $\mathcal{J}_\varphi(G)$.

Sparse graph classes with fpt model-checking



Sparse graph classes with fpt model-checking

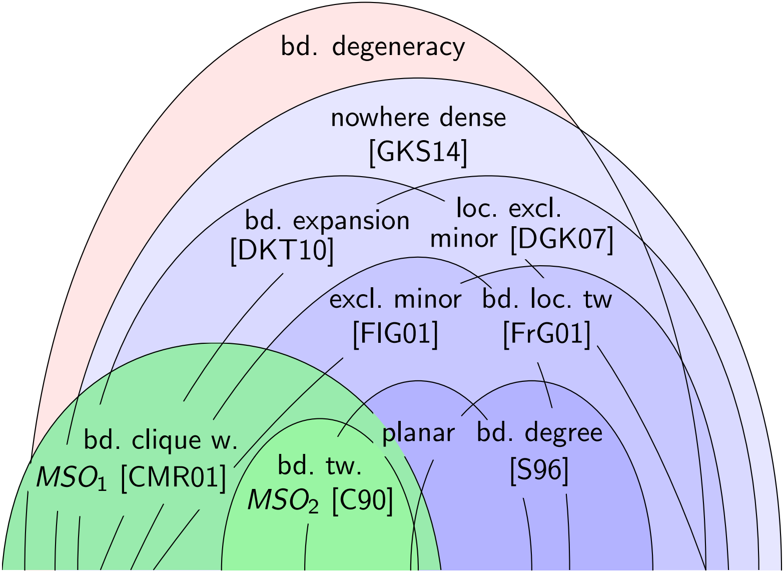


Model-checking beyond sparse graphs – Interpretations

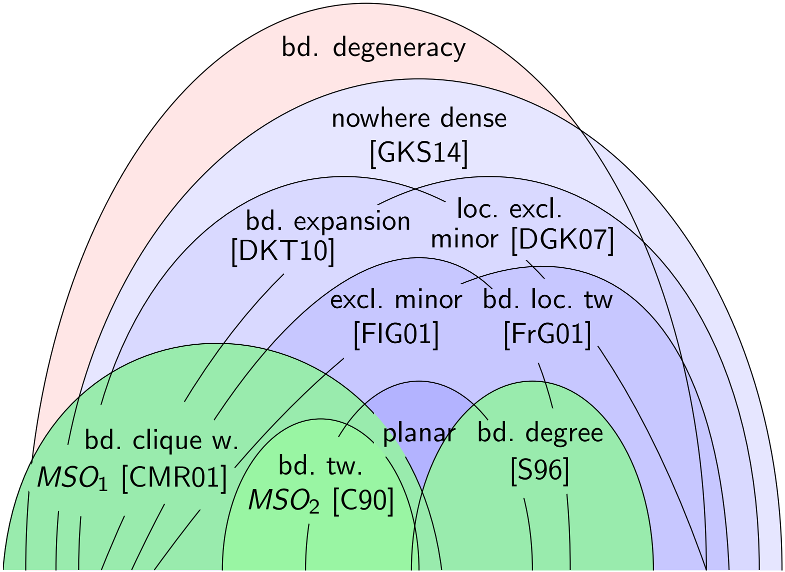
Theorem

- $\mathcal{C}$ has bounded clique-width $\iff$ $\mathcal{C}$ is an MSO_1-interpretation of a class of colored trees.
  $\Rightarrow$ an interpretation of a bounded clique-width graph has again bounded clique-width.
- A clique-decomposition of the input graph can be computed/approximated efficiently.
- MSO_1/FO model-checking on interpretations of bounded clique-width classes is fpt.

Sparse graph classes with fpt model-checking



Sparse graph classes with fpt model-checking



Model-checking beyond sparse graphs – Interpretations

Definition

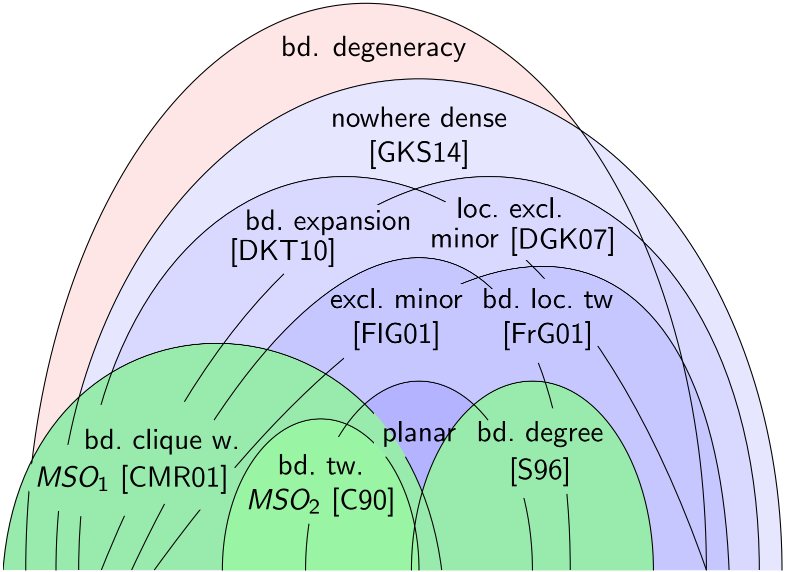
A class $\mathcal{C}$ of graphs is *near uniform* with parameter k if it satisfies certain conditions on the the *near-k-twin relation* defined by

$$u \sim_k v \iff |N(u) \triangle N(v)| \leq k.$$

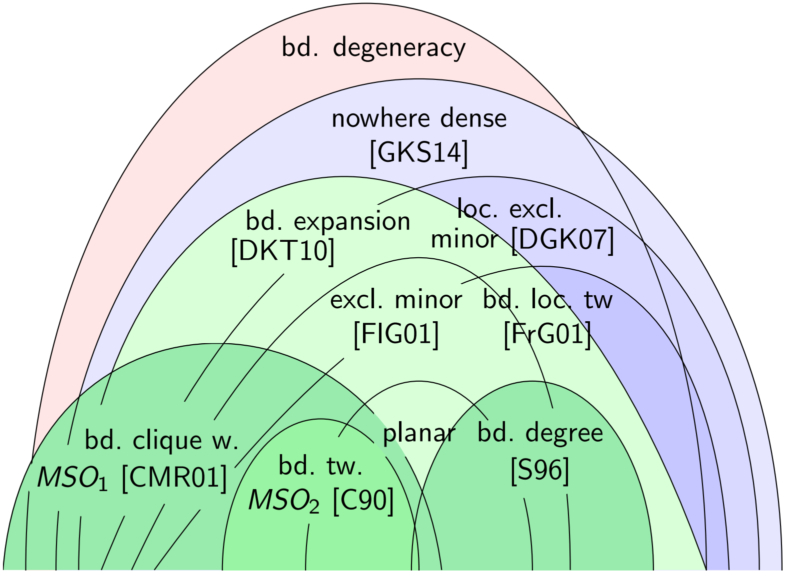
Theorem [Gajarský et al. 16]

- $\mathcal{C}$ near-uniform $\iff$ $\mathcal{C}$ interpretation of a bounded degree class.
- A bounded degree pre-image and an interpretation producing the input graph can be computed efficiently from a near uniform input graph.
- Consequently, model-checking on near uniform graphs is fpt.

Sparse graph classes with fpt model-checking



Work in progress



Open problem

Open problem

- Let $\mathcal{C}$ be nowhere dense,
- $\varphi(x, y) \in \mathrm{FO}$, and
- $H = \mathcal{I}_{\varphi}(G)$.
- Does there exist for every $p \in \mathbb{N}$ and every $\varepsilon > 0$ a coloring of $V(H)$ with $g(p, \varepsilon) \cdot n^{\varepsilon}$ colors such that the subgraph of H induced by any p color classes has clique-width/shrub-depth at most $f(p)$?

Model-checking beyond sparse graphs – Model-theory

Question

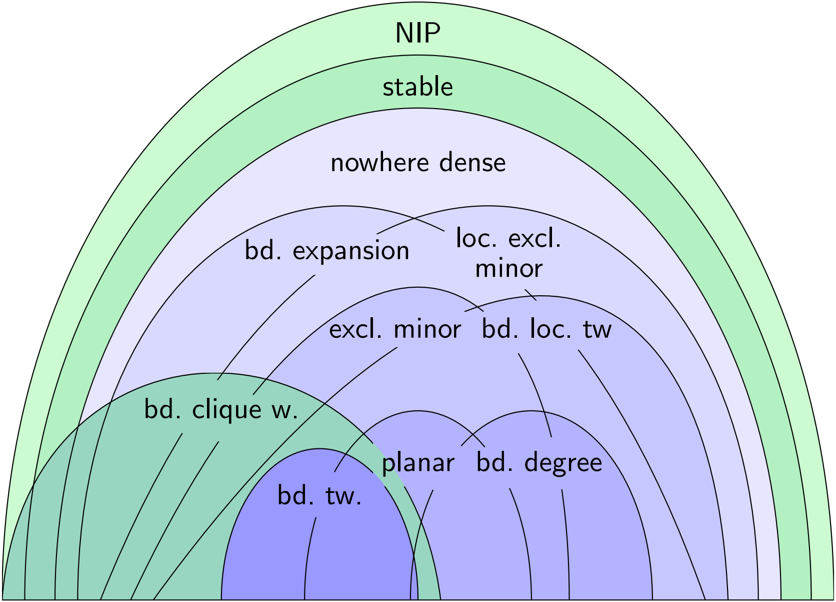
Could tractability of FO on nowhere dense classes be merely an artifact of tractability of FO on a much larger class which happens to coincide with nowhere dense classes if closed under subgraphs?

Theorem [Adler, Adler 14]

Let $\mathcal{C}$ be a class of graphs which is closed under taking subgraphs. The following are equivalent.

- $\mathcal{C}$ is nowhere dense.
- $\mathcal{C}$ is stable.
- $\mathcal{C}$ does not have the independence property (is NIP).

The model-theoretic notions



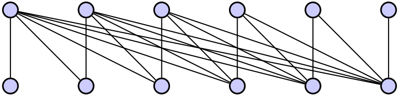
Stability

Research program

There is a huge number of combinatorial results which wait to be carried from infinite model theory to the finite.

Definition

A class $\mathcal{C}$ of graphs is *stable* if for every first-order formula $\varphi(\bar{x}, \bar{y})$ there is a constant c such that the interpretation of $G \in \mathcal{C}$ by φ excludes a *ladder* of length c as an induced subgraph.



Stability and Ramsey-type problems

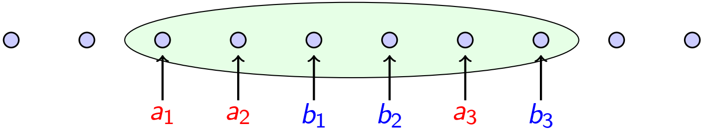
Definition

Let $(X,<)$ be a linearly ordered set. Let

$$[(X,<)]^k = \{(x_1,\dots,x_k) \in X^k, x_i < x_j \text{ for } i < j\}.$$

Let $f : [(X,<)]^k \to Z$.

A subordering $(Y,<)$ is *f-indiscernible* if f is constant on $[(Y,<)]^k$, i.e., $f(\bar{a}) = f(\bar{b})$ for all increasing k-tuples from Y.



$$f(a_1,a_2,a_3) = f(b_1,b_2,b_3)$$

Stability and Ramsey-type problems

Example

Let $G = (V, E)$ be a graph, let $<$ be an arbitrary order of V and let

$$f_E : [(V,<)]^2 \to \{0,1\} : uv \mapsto \begin{cases} 1 & \text{if } uv \in E(G) \\ 0 & \text{otherwise.} \end{cases}.$$

$(Y,<)$ is f_E-indiscernible if Y induces an edgeless graph or a complete graph in G.

Stability and Ramsey-type problems

Finite Ramsey Theorem

For all $k, m, n \in \mathbb{N}$ there exists $\ell \in \mathbb{N}$ such that

$$\ell \to (m)^k_n,$$

i.e., for every $(X, <)$ of cardinality ℓ and $f : [(X, <)]^k \to \{1, \dots, n\}$ there exists a subordering $(Y, <)$ of cardinality m which is f-indiscernible.

Example (continued)

Number ℓ such that $\ell \to (m)^2_2$ (as required for the function f_E) satisfies

$$(1 + o(1)) \frac{m}{\sqrt{2}e} 2^{m/2} \leq \ell \leq (1 + o(1)) \frac{4^{m-1}}{\sqrt{\pi} m}.$$

Stability and Ramsey-type problems

Definition

Let $\mathfrak{A}$ be a structure with universe A, $<$ a linear order on A and let Φ be a set of formulas. A subordering $(Y,<)$ of $(A,<)$ is Φ-*indiscernible* if for every $\varphi(x_1,\dots,x_k) \in \Phi$ and every pair $\bar{a},\bar{b} \in [(A,<)]^k$ we have

$$\mathfrak{A} \models \varphi(\bar{a}) \Longleftrightarrow \mathfrak{A} \models \varphi(\bar{b}).$$

We write

$$\ell \to (m)_\Phi,$$

if for every $(X,<)$ of cardinality ℓ there exists a subordering $(Y,<)$ of cardinality m which is Φ-indiscernible.

Theorem

If $\mathcal{C}$ is stable, then for all finite $\Phi \subseteq \mathrm{FO}$ there exists $t \in \mathbb{N}$ such that ℓ with $\ell \to (m)_\Phi$ satisfies $\ell \leq m^t$.

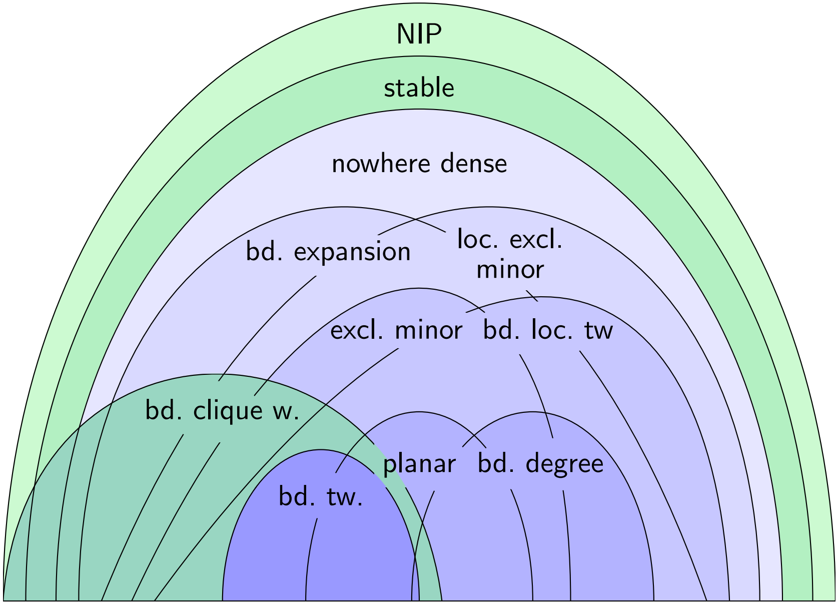
Algorithmic results based on stability

- We have not been able to apply these facts algorithmically on stable graphs, however,
- we have results for nowhere dense and $K_{i,j}$-free graph classes using stability related methods:

Exemplaric results

- Let $\mathcal{C}$ be nowhere dense. Then for every $r \in \mathbb{N}$ the distance-r dominating set problem admits a polynomial kernel on $\mathcal{C}$ [Kreutzer, Rabinovich, S. 16] and in fact an almost linear kernel [Eickmeyer et al. 17].
- Let $\mathcal{C} = \{G : K_{i,j} \nsubseteq G\}$. Then the dominating set problem is fixed-parameter tractable on $\mathcal{C}$ (a super simple proof for a result of Philip et al. 12).

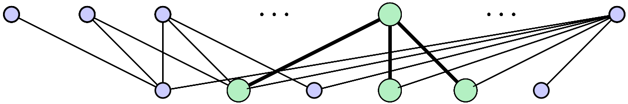
The model-theoretic notions



NIP

Definition

A class $\mathcal{C}$ of graphs is *NIP* if for every first-order formula $\varphi(\overline{x}, \overline{y})$ there is a constant c such that the interpretation of $G \in \mathcal{C}$ by φ has VC-dimension bounded by c.

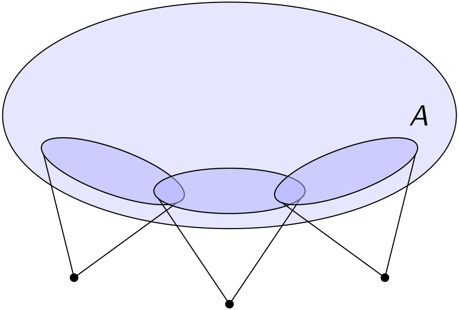


Methods for graphs of bounded VC-dimension

Sauer-Shelah Lemma (for graphs)

If G has VC-dimension c, then for all $A \subseteq V(G)$

$$|\{N(v) \cap A : v \in V(G)\}| \leq |A|^c.$$



Conclusion

- Many computational problems can be described elegantly in logics.
- In general, the model-checking problem for first-order logic is intractable.
- We search for the most general graph classes on which it is tractable.
- The classification for sparse (subgraph closed) classes is complete.
- The methods of interpretations is a strong method to generalize model-checking results. How far can we get?
- The notions of stability and NIP may be very useful in this context.
    - Remark: NIP classes are not the limit...