Model Theory of Tame Classes of Finite Structures

Part 2: Sparse Tame Classes

Anuj Dawar

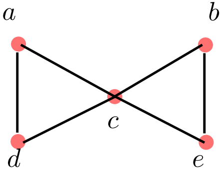
Department of Computer Science and Technology, University of Cambridge

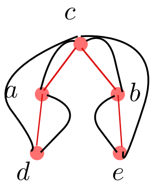
(with assistance from Ioannis Eleftheriadis)

Logic Colloquium, Gothenburg 24 June 2024

Treedepth

The *treedepth* *of a graph* G *is defined to be the smallest* k *such that the vertices of* G *can be arranged in a* *forest* *of height* k *so that every edge of* G *connects a vertex with an ancestor.*





We call such a forest an *elimination forest*.

Paths

Let G be a path that does not contain a path of length k.

Then, we can get an *elimination forest* of G by means of a *depth-first search*. The height of the forest is at most k.

Note: K_k, the *clique* on k vertices has paths of length $k-1$ and *treedepth* exactly k, so this bound is tight.

Note also that *treedepth* is *monotone* on the subgraph relation:
If G is a subgraph of H, then $\mathrm{td}(G) \leq \mathrm{td}(H)$.

Recursive Definition

The *treedepth* $\mathrm{td}(G)$ of a graph G can be characterized as follows:

$$\mathrm{td}(G) = \begin{cases} 1, & \text{if } |G| = 1; \\ 1 + \min_{v \in V(G)} \mathrm{td}(G \setminus v), & \text{if } G \text{ is connected and } |G| > 1; \\ \max_H \mathrm{td}(H) & \text{otherwise.} \end{cases}$$

where the *maximum* is taken over all *connected subgraphs* H of G.

Paths Again

Let P_k denote the graph consisting of a simple path of length k.

$$\mathrm{td}(P_k) = 1 + \min_{i \in [k-2]} (\mathrm{td}(P_i), \mathrm{td}(P_{k-1-i})) = 1 + \mathrm{td}(P_{\lceil n/2 \rceil})$$

Thus, $\mathrm{td}(P_k) = \lceil \log(k+1) \rceil$.

Consequently, a class of graphs $\mathcal{C}$ has bounded treedepth *if, and only if,* there is a bound on the length of the paths in graphs in $\mathcal{C}$.

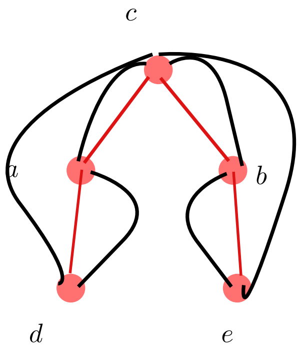
Splitter-Connector Game

Consider the following game played by two players we call *Splitter* and *Connector* on a (*possibly coloured*) graph G.

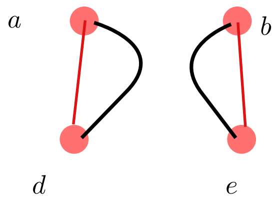
- $G_0 := G$;
- If G_i consists of a *single vertex*, *Splitter* has won;
- Otherwise: *Connector* chooses a *connected component* H of G and *Splitter* chooses a vertex $v \in H$.
- $G_{i+1} := H \setminus \{v\}$.

$\mathrm{td}(G) \leq k$ *if, and only if,* there is a *strategy* for *Splitter* to win in no more than k rounds.

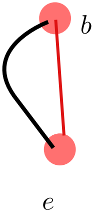
Elimination Forests and Splitter Strategies



Elimination Forests and Splitter Strategies



Elimination Forests and Splitter Strategies

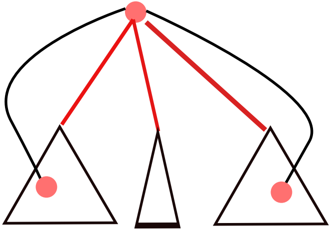


Well-Quasi Order

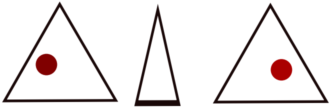
Classes of finitely coloured graphs of bounded treedepth are well-quasi ordered by the induced substructure relation.

Proof (sketch)
Prove by induction on k that, for any $s \in \mathbb{N}$: the class $\mathcal{T}_k(s)$ of s-coloured graphs of treedepth at most k is well-quasi ordered.
Base case follows from *Dickson's lemma*.
Inductively, let $(G_i)_{i \in I}$ be a sequence of graphs in $\mathcal{T}_{k+1}(s)$
Let H_i be a graph in $\mathcal{T}_k(s+1)$ obtained by deleting a vertex v from each component of G_i and colouring all its neighbours with a new colour.
By inductive hypothesis, there are $i < j$ with H_i an induced subgraph of H_j.
The result follows by our choice of colouring.

Elimination



Elimination



Extension Preservation

The extension preservation theorem holds in any class $\mathcal{C}$ of coloured graphs of bounded treedepth.

The *well-quasi ordering* guarantees that any sentence preserved under extensions has only finitely many *minimal models*.

The result extends easily to other vocabularies.

First-Order Evaluation

The problem of determining whether a first-order sentence holds in a graph is FPT on classes of graphs of bounded treedepth.

More precisely:

There is an algorithm, which takes a graph G and a sentence φ and determines whether $G \models \varphi$ and runs in time $f(k, l)n^2$ where

- f *is some computable function;*
- k *is the treedepth of* G*;*
- n *is the number of vertices in* G*; and*
- l *is the length of* φ*.*

First-Order Evaluation

Suffices to prove it for *basic local sentences*:

$$\exists x_1 \cdots \exists x_s \left(\bigwedge_{i \neq j} \delta(x_i, x_j) > 2r \wedge \bigwedge_i \psi^r(x_i) \right)$$

- Given G and φ, construct a (not necessarily optimal) elimination forest for G. This takes time $O(n^2)$.
- For each vertex $v \in G$, determine whether it satisfies the local condition $\psi^r(x_i)$. Aim to do this in time $O(n)$.
- Determine whether there is a $2r$-scattered set of size s among the set of vertices that do.

Local Condition

Suppose G is connected and let v be the *root* of its *elimination tree*.
Let H be the coloured graph obtained by deleting v and *colouring* every neighbour of v with a new colour.

We can compute from ψ^r a sentence χ_1 and a formula $\chi_2(x)$ in the expanded vocabulary such that:

- $H \models \chi_1$ *if, and only if,* $G \models \psi^r[v]$.
- $H \models \chi_2[u]$ *if, and only if,* $G \models \psi^r[u]$.

Recursively evaluating these formulae in H then suffices.

Nowhere Dense Classes

Nowhere-dense classes can be characterized by a *local* version of the *recursive* definition of treedepth.

$$\mathrm{rank}_r(G) = \begin{cases} 1, & \text{if } |G| = 1; \\ 1 + \min \mathrm{rank}_r(G \setminus v), & \text{if } G \text{ has radius} \leq r \text{ and } |G| > 1; \\ \max \mathrm{rank}_r(B_r(G, v)) & \text{otherwise.} \end{cases}$$

where the $\min$ and $\max$ are taken over all *verrices* of G.

A class $\mathbb{C}$ is *nowhere dense* if there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that for every $r \in \mathbb{N}$ and every $G \in \mathbb{C}$ it holds that $\mathrm{rank}_r^S(G) \leq f(r)$.

Splitter-Connector Game

A corresponding *Splitter-Connector* Game.

- $G_0 := G$;
- If G_i contains *no edges*, *Splitter* has won;
- Otherwise: *Connector* chooses a *vertex* u of G_i and *Splitter* chooses a vertex $v \in N_r^{G_i}(u)$.
- $G_{i+1} := N_r^{G_i}(u) \setminus \{v\}$.

$\mathrm{rank}_r(G) \leq k$ *if, and only if,* there is a *strategy* for *Splitter* to win in no more than k rounds.

$\mathrm{td}(G) = \mathrm{rank}_\infty(G)$.

Examples

For *any* graph G and any r , $\mathrm{rank}_r(G) \leq \mathrm{td}(G)$.
Thus, any class of bounded treedepth is *nowhere dense*.

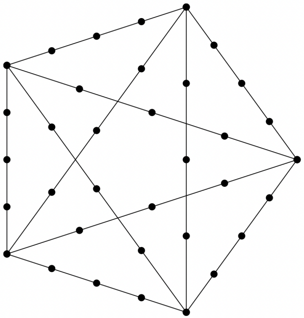
If G has *degree* at most d, for every vertex v: $|N_r(v)| \leq d^r$, so $\mathrm{rank}_r(G) \leq d^r$.

For any *tree* T, $\mathrm{rank}_r(T) \leq r + 1$.

If G is planar, $\mathrm{rank}_r(G) \leq 3r^3 + 5r + 2$.

$\mathrm{rank}_r(K_n^r) = n$ where K_n^r is the r-subdivided n-clique.

Subdivided Cliques



K_5^3

If a class $\mathcal{C}$ contains arbitrarily large K_n^r for some fixed r, it is *not* nowhere-dense.

Evaluating First-Order Formulas

Theorem (Grohe, Kreutzer, Siebertz)
For any *nowhere dense* class of graphs $\mathcal{C}$ and any $\epsilon > 0$, there is an algorithm deciding $G \models \varphi$ in time $O(f(l)n^{1+\epsilon})$ for some function f.

Suffices to prove it for *basic local sentences*:

$$\exists x_1 \cdots \exists x_s \left(\bigwedge_{i \neq j} \delta(x_i, x_j) > 2r \wedge \bigwedge_i \psi^r(x_i) \right)$$

We want to evaluate ψ^r at *every element*.

We aim to reproduce the recursive strategy that works with *treedepth*

Two main difficulties arise.

Recursion with Local Formulas

To determine $G \models \psi^r[v]$

Calculate a formula χ such that it suffices to check $\mathrm{Nbd}_r^{G-v}(u) \models \chi$ for all u.

The *quantifier rank* of χ can be exponential in r.
As we recurse, this means the *radius* r we need to consider increases.

The problem doesn't arise with *bounded treedepth* as we start with $r = \infty$

The solution is a version of *rank-preserving Gaifman locality*.

Number of Neighbourhoods

Calculate a formula χ such that it suffices to check $\mathrm{Nbd}_r^{G-v}(u) \models \chi$ for all u.

The neighbourhoods $\mathrm{Nbd}_r^{G-v}(u)$ are not pairwise disjoint.

The total number of neighbourhoods we need to consider in a recursion of depth l can be n^l.

But l depends on φ.

The solution is that we can *cover* the graph with a *small* number of sets of small *diameter* such that every r-neighbourhood is contained in such a set.

Neighbourhood Covers

For a graph $G = (V, E)$ and $r \in \mathbb{N}$, a family $\mathcal{X}$ of subsets of V is a *distance-r neighborhood cover* of G if for every $v \in V$, there is an $X \in \mathcal{X}$ with $N_r(v) \subseteq X$.

The *radius* of $\mathcal{X}$ is the largest *diameter* of any connected component in a graph induced by $X \in \mathcal{X}$.

The *overlap* of $\mathcal{X}$ is the largest m such that some $v \in V$ appears in m distinct sets in $\mathcal{X}$.

Example

A *small diameter* cover means, in our recursion we don't increase radius too much.

A *small overlap* means that the number of times each vertex is encountered in the course of a recursion is small, bounding overall *running time*

Given a graph $G \in \mathcal{BD}_d$, taking the set $\{N_r(v) \mid v \in G\}$ gives a distance r-neighbourhood cover of diameter $\leq 2r$ and overlap $\leq d^r$.

Neighbourhood Covers in Nowhere Dense Classes

For $\mathcal{C}$ a nowhere dense class of graphs, $r \in \mathbb{N}$ and an n-vertex graph $G \in \mathcal{C}$ and $\epsilon > 0$ there exists a distance-r neighborhood cover of G with diameter at most $4r$ and whose overlap at most $f(\mathcal{C}, r, \epsilon)(n^\epsilon)$.

Moreover this neighbourhood cover can be computed efficiently.

Theorem (Grohe, Kreutzer, Siebertz)
For any *nowhere dense* class of graphs $\mathcal{C}$ and any $\epsilon > 0$, there is an algorithm deciding $G \models \varphi$ in time $O(f(l)n^{1+\epsilon})$ for some function f.