Evaluating Formulas on Sparse Graphs

Part 3

Anuj Dawar

University of Cambridge Computer Laboratory

PhD Open, Warsaw, 5 November 2010

Review

We consider the complexity of the problem of deciding,

> Given a graph G and a formula φ
>
> whether $G \models \varphi$

when φ is either in FO or MSO.

In general the problem is PSPACE-complete and AW$[\star]$-hard.

We now to identify classes of *sparse* graphs where the problem becomes tractable.

Tractable here means *fixed-parameter tractable* with the formula length as parameter.

Results So Far

$\mathcal{T}_k$—the class of graphs of tree-width at most k.

$\mathcal{D}_k$—the class of graphs with maximal degree k.

Theorem (Courcelle)

For any MSO (or MS_2) sentence φ and any k there is a linear time algorithm that decides, given $G \in \mathcal{T}_k$ whether $G \models \varphi$.

Theorem (Seese)

For every sentence φ of FO and every k there is a linear time algorithm which, given a graph $G \in \mathcal{D}_k$ determines whether $G \models \varphi$.

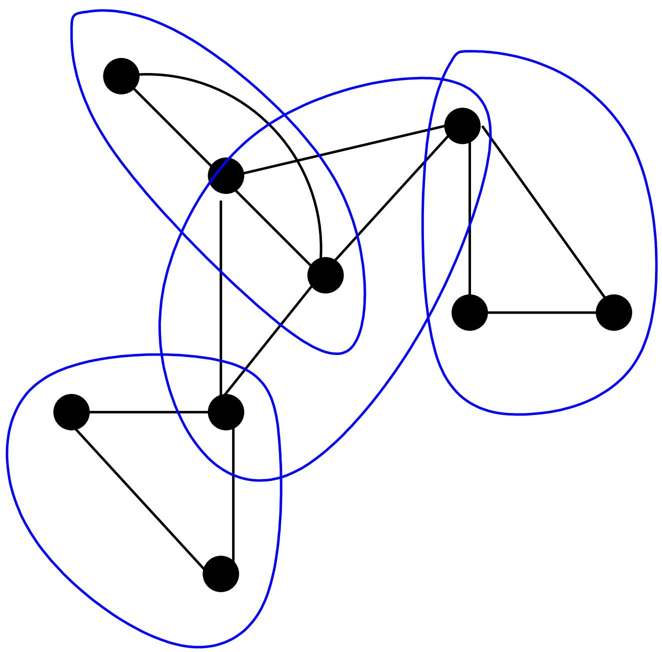
The proofs are based on two general methods:

- the *method of decompositions*; and
- the *method of locality*.

Treewidth

The *treewidth* of an undirected graph is a measure of how tree-like the graph is.

A graph has treewidth k if it can be covered by subgraphs of at most $k+1$ nodes in a tree-like fashion.



This gives a *tree decomposition* of the graph.

Treewidth

For a graph $G = (V, E)$, a *tree decomposition* of G is a relation $D \subset V \times T$ with a tree T such that:

- for each $v \in V$, the set $\{t \mid (v, t) \in D\}$ forms a connected subtree of T; and
- for each edge $(u, v) \in E$, there is a $t \in T$ such that $(u, t), (v, t) \in D$.

The *treewidth* of G is the least k such that there is a tree T and a tree decomposition $D \subset V \times T$ such that for each $t \in T$,

$$|\{v \in V \mid (v, t) \in D\}| \leq k + 1.$$

Courcelle's Theorem

Theorem (Courcelle)

For any MSO (or MS_2) sentence φ and any k there is a linear time algorithm that decides, given $G \in \mathcal{T}_k$ whether $G \models \varphi$.

Given $G \in \mathcal{T}_k$ and φ, compute:

- from G a labelled tree T; and
- from φ a bottom-up tree automaton $\mathcal{A}$

such that $\mathcal{A}$ accepts T if, and only if, $G \models \varphi$.

The Method of Decompositions

Suppose $\mathcal{C}$ is a class of graphs such that there is a finite class $\mathcal{B}$ and a finite collection Op of operations such that:

- $\mathcal{C}$ is contained in the closure of $\mathcal{B}$ under the operations in Op;
- there is a polynomial-time algorithm which computes, for any $G \in \mathcal{C}$, an Op-decomposition of G over $\mathcal{B}$; and
- for each m, the equivalence class $\equiv_m^{(\mathsf{MSO})}$ is an *effective* congruence with respect to to all operations $o \in$ Op (i.e., the $\equiv_m^{(\mathsf{MSO})}$-type of $o(G_1, \ldots, G_s)$ can be computed from the $\equiv_m^{(\mathsf{MSO})}$-types of $G_1, \ldots, G_s$).

Then, FO (MSO) satisfaction is fixed-parameter tractable on $\mathcal{C}$.

Relaxations of the Method

1. Instead of requiring $\mathcal{B}$ be finite, it sufficers to require that *satisfaction is in FPT over* $\mathcal{B}$.
2. In place of $\equiv_m^{(\text{MSO})}$, we can take any sequence of equivalence relations $\sim_m$ $(m \in \mathbb{N})$ satisfying
   - for every φ there is an m such that models of φ are closed under $\sim_m$; and
   - for all m, $\sim_m$ has finite index.

Bounded Degree Graphs

Theorem (Seese)

For every sentence φ of FO and every k there is a linear time algorithm which, given a graph $G \in \mathcal{D}_k$ determines whether $G \models \varphi$.

A proof is based on *locality* of first-order logic.

Note: this is not true for MSO unless P $=$ NP.

Gaifman's Locality Theorem

We write $\delta(x, y) > d$ for the formula of FO that says that the distance between x and y is greater than d.

We write $\psi^r(x)$ to denote the formula obtained from $\psi(x)$ by relativising all quantifiers to the set $N_r = \{y \mid \delta(x, y) < r\}$, i.e.

Each subformula $\exists y\theta$ is replaced by $\exists y(\delta(x, y) < r) \wedge \theta^r$

Each subformula $\forall y\theta$ is replaced by $\forall y(\delta(x, y) < r) \rightarrow \theta^r$

Gaifman's Locality Theorem

A *basic local sentence* is a sentence of the form

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

Theorem (Gaifman)

Every first-order sentence is equivalent to a Boolean combination of basic local sentences.

Seese's Theorem

How do we evaluate a basic local sentence
$\exists x_1 \cdots \exists x_s \left(\bigwedge_{i \neq j} \delta(x_i, x_j) > 2r \wedge \bigwedge_i \psi^r(x_i) \right)$ in a graph $G \in \mathcal{D}_k$?

For each $v \in G$, determine whether

$$N_r(a) \models \psi[a].$$

Since the size of $N_r(a)$ is bounded, this takes linear time.

Label a red if so. We now want to know whether there exists a $2r$-*scattered* set of red vertices of size s.

Finding a Scattered Set

(Frick and Grohe) describe a method to do this efficiently.

Choose red vertices from G in some order, removing the $2r$-neighbourhood of each chosen vertex.

$a_1 \in G,$
$a_2 \in G \setminus N_{2r}(a_1),$
$a_3 \in G \setminus (N_{2r}(a_1) \cup N_{2r}(a_2)), \ldots$

If the process continues for s steps, we have found a $2r$-scattered set of size s.

Otherwise, for some $u < s$ we have found $a_1, \ldots, a_u$ such that all red vertices are contained in

$$N_{2r}(a_1, \ldots, a_u)$$

This is a graph of bounded size and the property of containing a $2r$-scattered set of *red* vertices of size s can be stated in FO.

Method of Locality

- Suppose we have a computable function, associating a parameter $k_G \in \mathbb{N}$ with each graph G.
- Suppose we have an algorithm which, given G and φ decides $G \models \varphi$ in time

$$g(l, k_G)n^c$$

for some computable function g and some constant c.

- Let $\mathcal{C}$ be a class of graphs of *bounded local* k, i.e.

  there is a computable function $t : \mathbb{N} \to \mathbb{N}$ such that for every $G \in \mathcal{C}$ and $v \in G$, $k_{N_r(a)} < t(r)$.

Then, there is an algorithm which, given $G \in \mathcal{C}$ and φ decides whether $G \models \varphi$ in time

$$f(l)n^{c+1}$$

for some computable function f.

Planar Graphs

We now aim to combine the two methods to show the following

Theorem (Frick-Grohe)

For any $\varphi \in$ FO, there is a *quadratic* time algorithm that decides, given a *planar* graph G whether $G \models \varphi$.

The proof combines the methods of *decompositions* and *locality*.

Bounded Diameter Planar Graphs

The *diameter* of a graph G is the least d such that between any two vertices of G there is a path of length at most d.

The tractability of FO on planar graphs follows from the the following.

Theorem (Robertson-Seymour)

For every d there is a k such that any planar graph of diameter d has tree-width at most k.

Taking $k = 3d$ suffices. We sketch a proof of this.

Series Parallel Graphs

The class of *series-parallel graphs* consists of those graphs that can be obtained from a single edge: s—t

by operations of

- *series composition*

  This takes graphs (G_1, s_1, t_1) and (G_2, s_2, t_2) and gives the graph (G, s_1, t_2) that is formed by taking their disjoint union while identifying t_1 with s_2.

- *parallel composition*

  This takes graphs (G_1, s_1, t_1) and (G_2, s_2, t_2) and gives the graph (G, s, t) that is formed by taking their disjoint union while identifying s_1 with s_2 and t_1 with t_2..

This is *exactly* the class of graphs of tree-width 2.

Outerplanar Graphs

G is said to be *outerplanar* if it is *planar* and has a planar embedding in which all vertices are on the *outer face*.

Any outerplanar graph is a *series parallel* graph and therefore has treewidth at most 2.

Decomposing Planar Graphs

Suppose G is a *2-connected* planar graph of diameter d.

For graphs that are not 2-connected, we decompose the 2-connected components separately, as they are joined together in a tree-like fashion.

If G is outerplanar, we are done.

Otherwise, pick a planar embedding of G, an interior vertex v and two paths A and B to vertices u and w on the outer face.

$A \cup B$ is a set of at most $2d + 1$ vertices which separates the graph into G_1 and G_2.

Our aim is to show that each of $G_1 \cup A \cup B$ and $G_2 \cup A \cup B$ has a tree decomposition of width $3d$ in which $A \cup B$ appears inside a single bag.

Decomposing Planar Graphs

In G_1 (the situation for G_2 is symmetric), we choose another path C from v to a vertex between u and w on the outer face.

The decomposition of G_1 now consists of a bag containing $A \cup B \cup C$ and a decomposition (obtained *recursively*) of the two parts bounded by $A \cup B$ and $A \cup C$.

see picture.

The base cases are when the graph is outerplanar, or there is no vertex between u and w.

In the latter case, we recursively decompose the graph obtained by removing the edge (u, w).

Local Tree-Width

Let $t : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function.

LTW$_t$—the class of graphs G such that for every $v \in V(G)$:

$N_r^G(v)$ has tree-width at most $t(r)$. **(Eppstein; Frick-Grohe)**.

We say that $\mathcal{C}$ has *bounded local tree-width* if there is some function t such that $\mathcal{C} \subseteq$ LTW$_t$.

Examples:

1. $\mathcal{T}_k$ has local tree-width bounded by the constant function $t(r) = k$.
2. $\mathcal{D}_k$ has local tree-width bounded by $t(r) = k^r + 1$.
3. Planar graphs have local tree-width bounded by $t(r) = 3r$.

Bounded Local Tree-Width

Theorem (Frick-Grohe)

For any class $\mathcal{C}$ of bounded local tree-width and any $\varphi \in$ FO, there is a *quadratic* time algorithm that decides, given $G \in \mathcal{C}$, whether $G \models \varphi$.

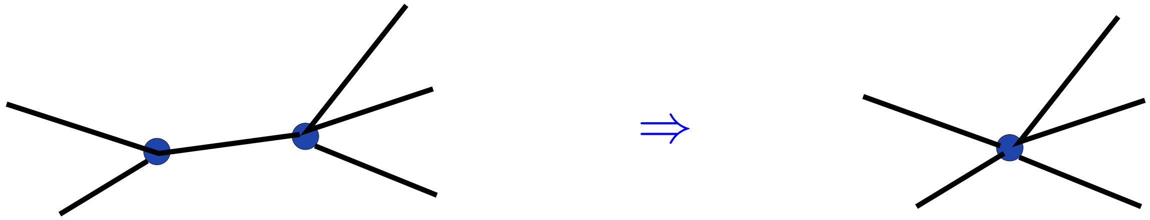
The proof is a direct application of the method of locality.

In place of planar graphs, we can take graphs embeddable in any *fixed surface* and obtain that FO satisfaction is fixed-parameter tractable as a consequence of the above.

Graph Minors

We say that a graph G is a minor of graph H (written $G \preceq H$) if G can be obtained from H by repeated applications of the operations:

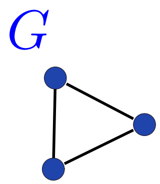
- *delete an edge*;
- *delete a vertex* (and all incident edges); and
- *contract an edge*

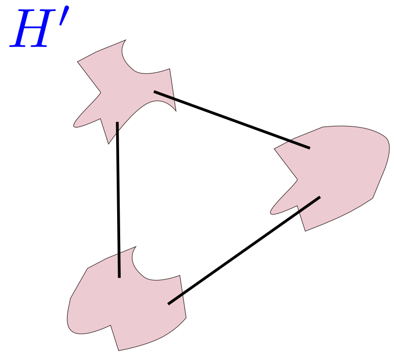


Graph Minors

Alternatively, $G = (V, E)$ is a minor of $H = (U, F)$, if there is a graph $H' = (U', F')$ with $U' \subseteq U$ and $F' \subseteq F$ and a surjective map $M : U' \to V$ such that

- for each $v \in V$, $M^{-1}(v)$ is a connected subgraph of H'; and
- for each edge $(u, v) \in E$, there is an edge in F' between some $x \in M^{-1}(u)$ and some $y \in M^{-1}(v)$.





Facts about Graph Minors

- G is planar if, and only if, $K_5 \not\preceq G$ and $K_{3,3} \not\preceq G$.
- If $G \subset H$ then $G \preceq H$.
- The relation $\preceq$ is transitive.
- If $G \preceq H$, then $\mathrm{tw}(G) \leq \mathrm{tw}(H)$.
- If $\mathrm{tw}(G) < k - 1$, then $K_k \not\preceq G$.

Say that a class of graphs $\mathcal{C}$ *excludes* H *as a minor* if $H \not\preceq G$ for all $G \in \mathcal{C}$.

$\mathcal{C}$ has *excluded minors* if it excludes some H as a minor (equivalently, it excludes some K_k as a minor).

- $\mathcal{T}_k$ excludes K_{k+2} as a minor.

More Facts about Graph Minors

Theorem (Robertson-Seymour)

In any infinite collection $\{G_i \mid i \in \omega\}$ of graphs, there are i, j with $G_i \preceq G_j$.

Corollary

For any class $\mathcal{C}$ *closed under minors*, there is a finite collection $\mathcal{F}$ of graphs such that $G \in \mathcal{C}$ *if, and only if,* $F \not\preceq G$ for all $F \in \mathcal{F}$.

Theorem (Robertson-Seymour)

For any G there is an $O(n^3)$ algorithm for deciding, given H, whether $G \preceq H$.

Corollary

Any class $\mathcal{C}$ closed under minors is decidable in *cubic time*.

Excluded Minor Classes

Write $\mathcal{M}_k$ for the class of graphs G such that $K_k \not\preccurlyeq G$.

Theorem (Flum-Grohe)

For $G \in \mathcal{M}_k$, $G \models \varphi$ is decidable in time $f(\varphi)n^5$.

We sketch some of the ideas behind the proof.

Decomposing Graphs with Excluded Minors

Robertson and Seymour show how to obtain a decomposition of graphs in $\mathcal{M}_k$.

Grohe shows that this can be done over graphs of *almost bounded local tree-width*.

Let

$$\mathcal{L}_\lambda = \{G \mid \forall H \prec G : \ \mathrm{ltw}_r(H) \leq \lambda r\}$$

$$\mathcal{L}_{\lambda,\mu} = \{G \mid \exists v_1, \ldots, v_\mu : \ G \setminus \{v_1, \ldots, v_\mu\} \in \mathcal{L}_\lambda\}$$

Almost Bounded Local Tree-width

Classes $\mathcal{L}_\lambda$ and $\mathcal{L}_{\lambda,\mu}$ are *minor-closed* and so decidable in cubic time.

Given $G \in \mathcal{L}_{\lambda,\mu}$, we can find $v_1, \ldots, v_\mu$ witnessing this in time $O(n^4)$.

For each v, check if $G - v$ is in $\mathcal{L}_{\lambda,\mu-1}$.

If so, add v to the list and proceed with $G - v$ and $\mathcal{L}_{\lambda,\mu-1}$.

Question: Is this algorithm in time $O(f(\lambda,\mu)n^4)$ for a *computable* function f?

There is a polynomial-time computable map taking a $G \in \mathcal{L}_{\lambda,\mu}$ to a *coloured graph* $G' \in \mathcal{L}_\lambda$ so that the FO-type of G is determined by that of G'.

G' is obtained from $G \setminus \{v_1, \ldots, v_\mu\}$ by adding new relations $S_1, \ldots, S_\mu$ interpreted by the neighbours of $v_1, \ldots, v_\mu$.

Decomposition Theorem

$\forall k \exists \lambda \exists \mu$

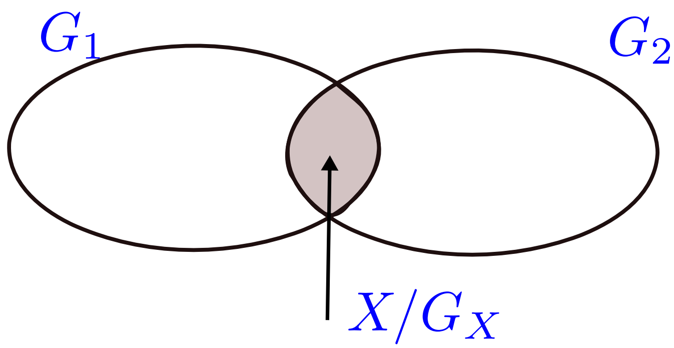
Any $G \in \mathcal{M}_k$ can be obtained from graphs in $\mathcal{L}_{\lambda,\mu}$ by a finite sequence of *clique sum* operations.

And the decomposition can be computed in time $O(n^4)$

Clique Sum: G_1, G_2 graphs with $X \subseteq G_1 \cap G_2$ a set of vertices that induces a clique in each of G_1 and G_2.

$$G_1 \oplus_{X,G_X} G_2$$

Take the disjoint sum of G_1 and G_2, identifying the two copies of X and replacing the clique by the graph G_X.



Congruences

For graphs $G \in \mathcal{L}_{\lambda,\mu}$, if X is a clique in G,

$$|X| < \lambda + \mu + 1$$

Thus, there are only finitely many operations of the form $\oplus_{X,G_X}$.

We have nearly satisfied the requirements for an application of the *automata-theoretic method*, but . . . .

If $X = x_1, \ldots, x_s$, the $\equiv_m$-type of $(G, x_1, \ldots, x_s)$, where

$$G = G_1 \oplus_{X,G_X} G_2,$$

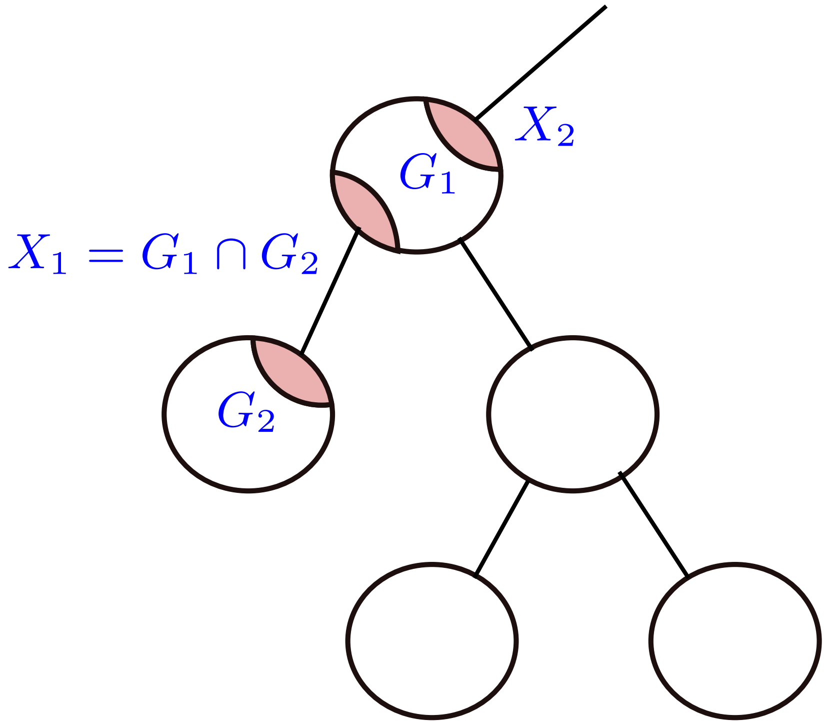
is given by the $\equiv_m$-types of $(G_1, x_1, \ldots, x_s)$ and $(G_2, x_1, \ldots, x_s)$.

However, different clique-sum operations may apply to different cliques X.

Bounding decompositions

While in a *bounded-width* tree-decomposition of G, the size of the individual bags is bounded, here we only have a bound on the size of the intersections between bags.

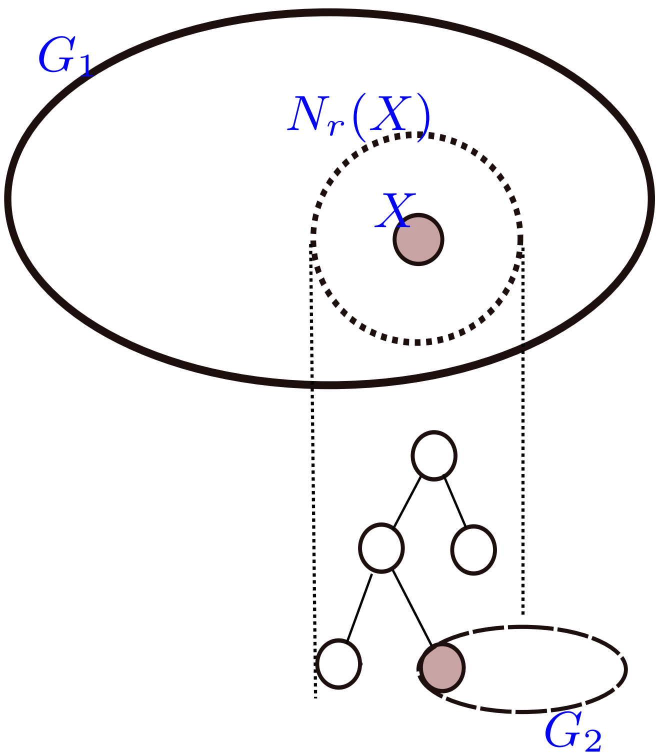
What we do have is a bound on the *local tree-width* of the bags G_1 (by replacing graphs in $\mathcal{L}_{\lambda,\mu}$ by their coloured companions in $\mathcal{L}_\lambda$).



Idea: the type of X_2 in $G_1 \oplus_X G_2$ is determined by the type of $(G_1, \bar{x_2})$, the type of $(G_2, \bar{x_1})$ and the *local neighbourhood* of the clique X_1 in G_1.

Typing the Sum

The tree-decomposition of $N_r^{G_1}(X)$ determines a function θ that takes the $\equiv_m$-type of $(G_2, \bar{x_2})$ to the $\equiv_m$-type of $N_r^{G_1}(X) \oplus_X (G_2, \bar{x_2})$



There are only finitely many such functions θ.

Define the asymmetric clique-sum *of type* θ:

$$(G_1, \bar{y}) \oplus^{\theta}_{X,G_X} (G_2, \bar{x})$$

of taking the clique-sum of the two graphs, joining $\bar{x}$ to a clique in G_1 whose neighbourhood has type θ.

Automata on $\mathcal{M}_k$

Given a first-order sentence φ, it determines a radius of locality r and quantifier rank m.

- We have a finite collection of operations $\oplus^{\theta}_{X,G_X}$ (depending on r and m).
- We have structures $(G, \bar{x})$, where the length of x is bounded by s (depending only on k).

  Thus, there are only finitely many $\equiv_m$ classes.

- $\equiv_m$ is a congruence for each operation $\oplus^{\theta}_{X,G_X}$.

Thus, satisfaction for *first-order logic* is *fixed-parameter tractable* on $\mathcal{M}_k$.

(Flum-Grohe)

Results So Far

1. $\mathcal{T}_k$—the class of structures of tree-width at most k.

   Courcelle (1990) shows that every MSO definable property is decidable in linear time on this class.

2. $\mathcal{D}_k$—the class of structures of *degree* bounded by k.

   Seese (1996) shows that every FO definable property is decidable in linear time.

3. LTW$_t$—the class of structures of *local tree-width* bounded by a function t.

   Frick and Grohe (2001) show that every FO definable property is decidable in quadratic time.

4. $\mathcal{M}_k$—the class of structures *excluding* K_k *as a minor*.

   Flum and Grohe (2001) show that every FO definable property is decidable in time $O(n^5)$.

Map of Classes