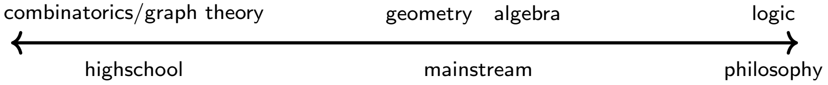
Understanding Some Graph Parameters by Infinite Model Theory

Yijia Chen
Fudan University

November 14, 2020 at Nankai

Joint work with Jörg Flum (Freiburg)

Spectrum of mathematics



Theorem (Glebskij, Kogan, Liogon'kij, Talanov, 1969; Fagin, 1976)

*Every graph property definable in first-order logic (*FO*) is almost surely true or almost surely false.*

Corollary

Parity cannot be defined by FO*.*

Theorem (Courcelle, 1990)

*Every graph property definable in monadic second-order logic (*MSO*) can be decided in linear time on graphs of bounded tree-width.*

Corollary

The 3-colorability problem can be decided in linear time on graphs of bounded tree-width.

In most cases, logic provides a general framework encompassing a large number of concrete problems, and graph theory provides proof methods and algorithmic tools.

Theorem (Ding, 1992)

Let $k \geq 1$. Then there are finitely many graphs $H_1, \ldots, H_{m_k}$ such that for any graph G

$$G \text{ has a tree-depth at most } k \iff \text{no } H_i \text{ is an induced subgraph of } G.$$

Remark

Tree-depth is a parameter which measures how close a graph is to a star.

I will sketch a new proof by the Łoś-Tarski Theorem from model theory.

Why the new logic proof might be interesting?

1. The original proof is purely combinatorial, using Higman's Lemma on well quasi-ordering.
2. The result is originally about finite graphs, and the combinatorial proof does not apply to infinite graphs.
3. Our proof, thus the result, applies to infinite graphs, while our tool – the Łoś-Tarski Theorem fails in the finite.
4. By the Completeness Theorem, we can compute the forbidden subgraphs $H_1, \ldots, H_{m_k}$ from k by the logic proof. This is not directly implied by the combinatorial proof.

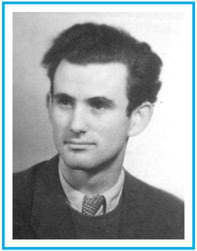
What is (infinite) model theory?

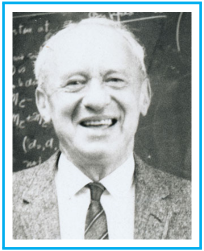
Model theory is the study of classes of mathematical structures (e.g. groups, fields, graphs, universes of set theory) from the perspective of mathematical logic.

Model theory recognizes and is intimately concerned with a duality: it examines semantical elements (meaning and truth) by means of syntactical elements (formulas and proofs) of a corresponding language.

–Wikipedia

Łoś-Tarski Theorem





Theorem

Let K be a graph property, equivalently K is a class of graphs (finite and infinite) that is closed under isomorphisms. Then the following are equivalent.

(i) K is definable in first-order logic and preserved under induced subgraphs, which is a semantic property.

(ii) K is definable by a universal FO-sentence, which is a syntactic property.

Why the emphasis on infinite?

Most powerful tools of logic deal with infinite structures, but many deep results in graph theory are about finite graphs.

It has been long argued that classical infinite model theory is not suitable for finite structures, thus for (finite) combinatorics and computer science finite model theory is the right model theory.

Plan

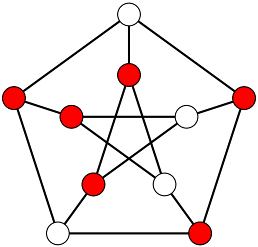
1. Vertex cover – forbidden induced subgraph characterization.
2. Tree-depth
    2.1 forbidden induced subgraph characterization
    2.2 MSO = FO, implying many NP-hard problems can be solved efficiently on graphs of small tree-depth.
3. Shrub-depth
    3.1 forbidden induced subgraph characterization
    3.2 MSO = FO.

Vertex Cover

Definition

Let $G = (V, E)$ be a graph. Then a vertex cover is a vertex subset $S \subseteq V$ such that for every $\{u, v\} \in E$

$$\{u, v\} \cap S \neq \emptyset.$$



1. The vertex cover problem is among the first problems discovered to be NP-complete.
2. There are many exact/approximation/parameterized algorithms for the vertex cover problem.
3. Vertex cover number has been studied as a graph parameter, i.e., the minimum number of vertices which need to be removed to make a graph edge-less.

László Lovász

Theorem

There are finitely many graphs $H_1, \ldots, H_{m_k}$ such that

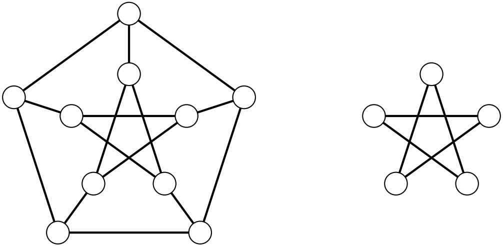
G *has a vertex cover of size at most* k

$\iff$ *no* H_i *is an induced subgraph of* G.

The Łoś-Tarski Theorem

Let G and H be two graphs (finite or infinite). G is an induced subgraph of H if $V(G) \subseteq V(H)$ and

$$E(G) = E(H) \cap V(G) \times V(G).$$



Definition

A first-order logic (FO) sentence φ is preserved under induced subgraphs if for any graphs G and H where G is an induced subgraph of H

$$H \models \varphi \quad \text{implies} \quad G \models \varphi$$

Example

For any $k \geq 1$,

$$\forall x_1 \cdots \forall x_{k+1} \bigvee_{1 \leqslant i < j \leqslant k+1} x_i = x_j$$

is preserved under induced subgraphs.

Definition

An FO-sentence φ is universal if

$$\varphi = \forall x_1 \cdots \forall x_k \psi,$$

where ψ is quantifier-free.

Theorem (trivial)

Any universal sentence is preserved under induced subgraphs.

Theorem

Let φ *be an* FO*-sentence which is preserved under induced subgraphs. Then there is a universal sentence* ψ *such that*

$$\models \varphi \leftrightarrow \psi.$$

The semantic property of the closure of induced subgraphs is equivalent to the syntactic property of being universal FO.

Theorem (Tait, 1959; *C.* and Flum, 2020)

There is an FO*-sentence which is preserved under induced subgraphs in finite, i.e., for any finite graphs G and H where G is an induced subgraph of H*

$$H \models \varphi \quad \textit{implies} \quad G \models \varphi,$$

such that φ is not equivalent to any universal sentence.

The proof of the Łoś-Tarski Theorem uses the Compactness Theorem, which does not hold in finite.

Let

$$\varphi_k := \exists x_1 \cdots \exists x_k \forall u \forall v \left(Euv \to \bigvee_{1 \leqslant i \leqslant k} (u = x_i \vee v = x_i) \right).$$

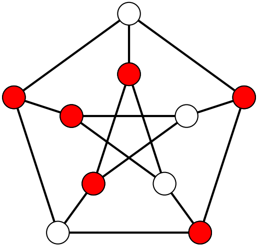
Then for any graph G

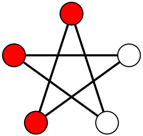
$$G \text{ has a vertex cover of size at most } k \quad \Longleftrightarrow \quad G \models \varphi_k.$$

φ_k is not universal.

Theorem (trivial)

Let G be a graph with a vertex cover of size at most k. Then any induced subgraph of G has a vertex cover of size at most k as well.





1. The k-vertex-cover problem can be defined by an FO-sentence φ_k (not universal).
2. The k-vertex-cover problem is closed under induced subgraphs.
3. Can we use the Łoś-Tarski Theorem?

No, graphs are finite graphs.

We allow graph $G = (V, E)$ to have infinite V, and infinite E as well.

Definition

Let $k \geq 1$ and $G = (V, E)$ be a graph (finite or infinite). Then a vertex cover is a vertex subset $S \subseteq V$ such that for every $\{u, v\} \in E$

$$\{u, v\} \cap S \neq \emptyset.$$

For any graph G, finite or infinite,

G has a vertex cover of size at most k

$$\iff G \models \exists x_1 \cdots \exists x_k \forall u \forall v \left(Euv \to \bigvee_{1 \leqslant i \leqslant k} \left(u = x_i \vee v = x_i \right) \right).$$

The k-vertex-cover problem on finite and infinite graphs is preserved under induced subgraphs.

Theorem (*C.* and Flum, 2020)

For any $k \geq 1$, there is a universal FO*-sentence ψ_k such that for any graph G*

$$G \text{ has a vertex cover of size at most } k \iff G \models \psi_k.$$

Compared to:

Theorem (Lovász)

There are finitely many graphs $H_1, \ldots, H_{m_k}$ such that

$$G \text{ has a vertex cover of size at most } k \iff \text{no } H_i \text{ is a subgraph of } G.$$

Lemma

For every universal sentence φ there are graphs $H_1, \ldots, H_s$ such that for any graph G

$$G \models \varphi \quad \iff \quad \text{no } H_i \text{ is an induced subgraph of } G$$

Proof.

1. φ can be written as
$$\bigwedge_{i \in I} \neg \exists x_1 \cdots \exists x_m \bigwedge_{j \in J_i} \gamma_{ij}.$$
2. For every $i \in I$ there are graphs $H_{i1}, \ldots, H_{i\ell_i}$ such that for any graph G
$$G \models \neg \exists x_1 \cdots \exists x_m \bigwedge_{j \in J_i} \gamma_{ij} \quad \iff \quad \text{no } H_{ij} \text{ is an induced subgraph of } G.$$

□

1. There is a universal sentence φ_k such that for any graph G

$$G \text{ has a vertex cover of size at most } k \quad \iff \quad G \models \psi_k.$$

2. There are graphs $H_1, \ldots, H_{s_k}$ such that for any graph G

$$G \text{ has a vertex cover of size at most } k$$
$$\iff \text{no } H_i \text{ is an induced subgraph of } G$$

1. Generalize vertex cover to infinite graphs.
2. The k-vertex-cover problem is definable in FO for finite and infinite graphs.
3. The k-vertex-cover problem is preserved under induced subgraphs.
4. By the Łos-Tarski Theorem there is a universal sentence which defines the k-vertex-cover problem.
5. Any characterization by a universal sentence is equivalent to a characterization of forbidden induced subgraphs.

Theorem (*C.* and Flum, 2020)

Let K *be a class of graphs (finite and infinite) such that*

1. K *is definable in* FO,
2. *for every* $G \in \mathsf{K}$ *and* H *an induced subgraph of* G, *then* $H \in \mathsf{K}$,

Then there are finitely many $H_1, \ldots, H_m$, *all finite graphs, such that for any graph* G

$$G \in \mathsf{K} \quad \Longleftrightarrow \quad \text{no } H_i \text{ is an induced subgraph of } G.$$

Tree-depth

Tree-depth

1. Tree-depth was introduced by J. Nešetřil and P. Ossona de Mendez in the theory of graphs of bounded expansion.
2. Tree-depth is equivalent to vertex ranking, ordered coloring, and elimination order.
3. Tree-depth measures how close a graph is to a star, similar as that tree-width measures how close a graph is to a tree.
4. Graphs of small tree-depth often admit fast parallel algorithms, similar as graphs of small tree-width admit fast sequential algorithms.

Definition

Let $G = (V, E)$ be a graph (finite). Then its tree-depth is

$$\mathsf{td}(G) := \begin{cases} 1 & \text{if } |V| = 1 \\ 1 + \min_{v \in V} \mathsf{td}(G \setminus v) & \text{if } V| \geq 2 \text{ and } G \text{ is connected} \\ \max_{\substack{C \text{ a connected} \\ \text{component of } G}} \mathsf{td}(C) & \text{if } G \text{ is not connected.} \end{cases}$$

Recall

Theorem (Ding, 1992)
Let $k \geq 1$. Then there are graphs $H_1, \ldots, H_{m_k}$ such that for any graph G

$$\mathsf{td}(G) \leqslant k \quad \Longleftrightarrow \quad \text{no } H_i \text{ is an induced subgraph of } G.$$

Theorem (*C.* and Flum, 2020)
Let K *be a class of graphs (finite and infinite) such that*

1. K *is definable in* FO,
2. *for every $G \in$ K and H an induced subgraph of G, then $H \in$ K,*

Then there are finitely many $H_1, \ldots, H_m$, all finite graphs, such that for any graph G

$$G \in \mathsf{K} \quad \Longleftrightarrow \quad \text{no } H_i \text{ is an induced subgraph of } G.$$

1. Generalize tree-depth to infinite graphs.
2. Define the class K of graphs G, finite and infinite, with $\mathrm{td}(G) \leqslant k$ in FO.
3. Prove that K is preserved under induced subgraphs.
4. Apply our meta-theorem.

We use exactly the same definition.

Definition

Let $G = (V, E)$ be a graph (finite or infinite). Then its tree-depth is

$$\mathsf{td}(G) := \begin{cases} 1 & \text{if } |V| = 1 \\ 1 + \min_{v \in V} \mathsf{td}(G \setminus v) & \text{if } V| \geq 2 \text{ and } G \text{ is connected} \\ \max_{\substack{C \text{ a connected} \\ \text{component of } G}} \mathsf{td}(C) & \text{if } G \text{ is not connected.} \end{cases}$$

There are infinite graphs G whose $\mathsf{td}(G)$ is not defined.

Let K be a class of graphs of bounded tree-depth.

Theorem (Elberfeld, Grohe, and T. Tantau, 2016)

*Every sentence in monadic second-order logic (*MSO*) is equivalent to an* FO*-sentence on* K.

Corollary

For graphs in K *the 3-colorability problem can be defined in* FO*, thus is decidable by* AC^0*-circuits, or equivalently solvable in constant parallel time.*

MSO is the restriction of second-order logic in which every second-order variable is a set variable.

A graph G is 3-colorable if and only if

$$G \models \exists X_1 \exists X_2 \exists X_3 \left(\forall u \bigvee_{1 \leqslant i \leqslant 3} X_i u \wedge \forall u \bigwedge_{1 \leqslant i < j \leqslant 3} \neg (X_i u \wedge X_j u) \right.$$
$$\left. \wedge \forall u \forall v \Big(Euv \rightarrow \bigwedge_{1 \leqslant i \leqslant 3} \neg (X_i u \wedge X_i v) \Big) \right).$$

MSO can also characterize SAT, INDEPENDENT-SET, DOMINATING-SET, etc.

On strings and trees the expressive power of MSO coincides with automata.

Let K be a class of graphs (finite or infinite) of bounded tree-depth.

Theorem

For any MSO*-sentence* φ *there is an* FO*-sentence* ψ *such that for any* $G \in \mathsf{K}$ *and an ordering* $<$ *on* G

$$G \models \varphi \iff (G, <) \models \psi.$$

The ordering $<$ is necessary to define a tree T of constant depth from G, and then we construct the FO-sentence to "simulate an automata corresponding to the MSO-sentence φ" on T.

Theorem

For any MSO*-sentence* φ *there is an* FO*-sentence* ψ *such that for any* $G \in \mathsf{K}$ *and an ordering* $<$ *on* G

$$G \models \varphi \iff (G, <) \models \psi.$$

Corollary

For any MSO*-sentence* φ *there are* FO*-sentences* $\psi(<_1)$ *and* $\psi(<_2)$ *such that for any graph* $G \in \mathsf{K}$ *and any two orderings* $<_1$ *and* $<_2$

$$G \models \varphi \iff (G, <_1) \models \psi(<_1) \iff (G, <_2) \models \psi(<_2).$$

In particular

$$(G, <_1, <_2) \models \psi(<_1) \to \psi(<_2)$$

Theorem

Let φ *and* ψ *be two* FO*-sentences with*

$$\models \varphi \to \psi. \tag{1}$$

Then there is an FO*-sentence* θ *which only contains symbols appearing both in* φ *and* ψ *such that*

$$\models \varphi \to \theta \quad \textit{and} \quad \models \theta \to \psi.$$

Remark

It is crucial for Craig's Interpolation Theorem that (1) *refers to all finite and infinite structures.*

Recall

Corollary

For any MSO*-sentence* φ *there are* FO*-sentences* $\psi(<_1)$ *and* $\psi(<_2)$ *such that for any* G*, finite or infinite, with* $\text{td}(G) \leqslant k$

$$(G, <_1, <_2) \models \psi(<_1) \to \psi(<_2)$$

By Craig's Interpolation Theorem and the FO definability of $\text{td}(G) \leqslant k$

Corollary

For any MSO*-formula* φ *there is an* FO*-sentence* ψ *(without* $<$*) such that for any graph* $G \in \mathsf{K}$

$$G \models \varphi \iff G \models \psi.$$

Shrub-depth

Shrub-depth is a graph parameter introduced by R. Ganian, P. Hliněný, J. Nesetril, J. Obdrzálek, P. Ossona de Mendez, and R. Ramadurai as the "tree-depth of dense graphs."

Lemma (trivial)

Let $G = (V, E)$ *be a (finite) graph with* $\mathsf{td}(G) \leqslant k$*. Then*

$$|E| \leqslant (k-1)(|V|-1).$$

That is, G *is sparse.*

A graph $G = (V, E)$ is sparse if $|E| = O(|V|)$.

Remark

Most graph parameters lead to sparse graphs, one notable exception is clique-width.

Clique-width can be viewed as "tree-width of dense graphs."

Shrubs are perennial woody plants, and therefore have persistent woody stems above ground (compare with herbaceous plants). Usually shrubs are distinguished from trees by their height and multiple stems.

– Wikipedia

dense shrub

sparse tree

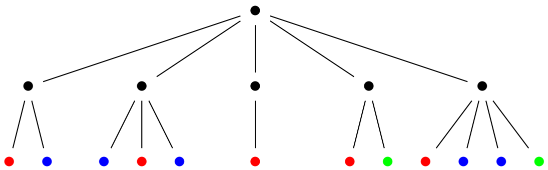
1. First we define the class $\textsc{Tree}[m, d]$ of trees of depth d with m labels.
2. Using a signature D any $T \in \textsc{Tree}[m, d]$ can be transferred to a shrub G.

$\textsc{Tree}[m, d]$ is the class of rooted trees with m labels and of depth d:

1. every root-to-leaf path has length d,
2. every leaf t is labelled with a color in $c(t) \in [m]$.

Example

A rooted tree in $\textsc{Tree}[3, 2]$.



Definition

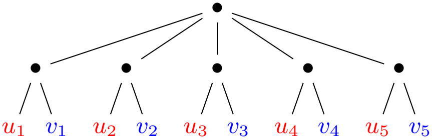
A tree-model of m labels and depth d of a graph G is a pair (T, D) of a rooted tree $T \in \textsc{Tree}[m, d]$ and a signature

$$D \subseteq \{1, 2, \ldots, m\}^2 \times \{2, 4, \ldots, 2 \cdot d\}$$

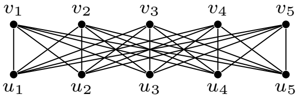
for some $h \geq d$ such that

1. for any $i, j \in [m]$ and $s \in [d]$ if $(i, j, s) \in D$, then $(j, i, s) \in D$,
2. $V(G) = \textit{leaves}(T)$,
3. $E(G) = \Big\{\{u, v\} \mid u, v \in V(G) \text{ and } (c(u), c(v), \textit{dist}^T(u, v)) \in D\Big\}$.

Examples

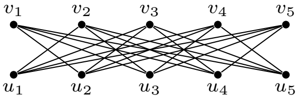


$$D := \{({\color{red}\bullet}, {\color{blue}\bullet}, 2), ({\color{red}\bullet}, {\color{blue}\bullet}, 4)\}$$



complete bipartite graph $K_{5,5}$

$$D := \{({\color{red}\bullet}, {\color{blue}\bullet}, 4)\}$$



almost complete bipartite graph $B_{5,5}$

Definition
$\mathrm{TM}_m(d)$ is the class of graphs with a tree-model in $\textsc{Tree}[m, d]$.

Definition
Let K be a class of graphs. Then K has shrub-depth d, if

1. $\mathsf{K} \subseteq \mathrm{TM}_m(d)$ for for some $m \in \mathbb{N}$,
2. $\mathsf{K} \not\subseteq \mathrm{TM}_{m'}(d-1)$ for for every $m' \in \mathbb{N}$.

Examples

1. The class of complete graphs has shrub-depth 1.
2. The class of complete bipartite graphs $K_{n,n}$ has shrub-depth 1.
3. The class of almost complete bipartite graphs $B_{n,n}$ has shrub-depth 2.
4. $\textsc{Tree}[1, d]$ has shrub-depth d.

Closure under induced subgraphs

Lemma (trivial)

Let $G \in \mathrm{TM}_m(d)$ *and* H *be an induced subgraph of* G*. Then* $H \in \mathrm{TM}_m(d)$*, too.*

Let $m, d \geq 1$.

Theorem (Gajarský et al., 2019)

There are graphs $H_1, \ldots, H_{\ell_{m,d}}$ such that for any graph G

$$G \in \mathrm{TM}_m(d) \iff \text{no } H_i \text{ is an induced subgraph of } G.$$

Theorem (Gajarský and Hliněný, 2015)

MSO = FO *on* $\mathrm{TM}_m(d)$.

1. Generalize $\textsc{Tree}[m, d]$ and $\mathrm{TM}_m(d)$ to infinite trees and infinite graphs.
2. We apply our meta-theorem to prove the forbidden induced subgraph characterization of $\mathrm{TM}_m(d)$. But this requires to show $\mathrm{TM}_m(d)$ is definable in FO.
3. We show that MSO $=$ FO on $\mathrm{TM}_m(d)$. A key step of the proof is that $\mathrm{TM}_m(d)$ can be defined approximately in FO.
4. We prove that $\mathrm{TM}_m(d)$ is definable in MSO. Then combined with 3, we conclude that $\mathrm{TM}_m(d)$ can be defined exactly in FO.

Conclusions

1. It has been long argued that classical infinite model theory is not suitable for finite structures, thus for (finite) combinatorics and computer science finite model theory is the right model theory.

   Our findings show that this might not be one hundred percent correct.

2. Our logic proofs of those combinatorial results use much less combinatorics than the original proofs.

   Pros: Reveal a certain "meta-structure" of those problems, and yield computable characterizations by the completeness theorem.

   Cons: Provide no explicit combinatorial constructions, the bounds of those characterizations directly derived from logic are very bad, sometimes even not recursive [*C.* and Flum, 2020].

Thank You!