Lovász meets Łoś and Tarski
– understanding forbidden induced subgraphs by model theory

Yijia Chen
Shanghai Jiao Tong University

Joint work with Jörg Flum (Freiburg)

May 13th, 2021, SCMS

K. B. R. Kolipaka and M. Szegedy

Moser and Tardos meet Lovász

In *Proceedings of the Proceedings of the 43rd annual ACM symposium on Theory of Computing*, 235-244, 2011.

H. Dell, M. Grohe, and G. Rattan

Lovász meets Weisfeiler and Leman

In *Proceedings of the 45th International Colloquium on Automata, Languages, and Programming*, 40:1-40:16, 2018.

Moser and Tardos meet Lovász

Algorithms meet probability.

Lovász meets Weisfeiler and Leman

Combinatorics meets algorithms.

Lovász meets Łoś and Tarski

Combinatorics meets logic.

Model theory can give logic proofs of theorems in pure combinatorics.

Logic might help us to understand why some combinatorial problems are hard.

László Lovász

Theorem

For every $k \geq 1$ there are finitely many graphs $H_1, \ldots, H_{m_k}$ such that

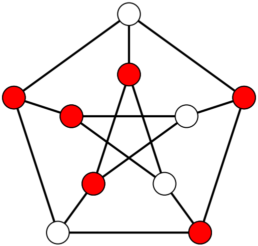
G has a vertex cover of size at most k

$\iff$ *no H_i is a subgraph of G.*

Definition

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

$$u \in S \quad \text{or} \quad v \in S.$$



Fix $k \geq 1$. The k-vertex-cover problem is

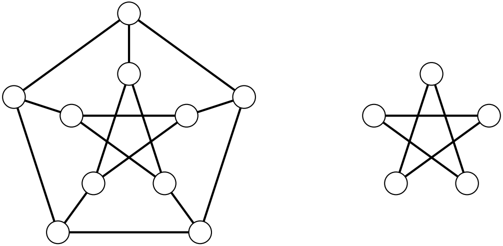
Input:	A graph $G = (V, E)$.
Problem:	Decide whether G has a vertex cover of size at most k.

Alternatively, you might view it as the class of graphs G with a vertex cover of size at most k. Then Lovász's result says that it can be characterized by a finite set of forbidden subgraphs.

The Łoś-Tarski Theorem

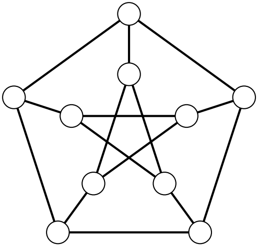
Let $\mathcal{A}$ and $\mathcal{B}$ be two relational structures (finite or infinite). $\mathcal{A}$ is an induced substructure of $\mathcal{B}$ if $A \subseteq B$ and if for every relation symbol R,

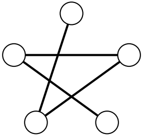
$$R^{\mathcal{A}} = R^{\mathcal{B}} \restriction A.$$



$\mathcal{A}$ is a substructure of $\mathcal{B}$ if $A \subseteq B$ and if for every relation symbol R,

$$R^{\mathcal{A}} \subseteq R^{\mathcal{B}} \upharpoonright A.$$





Definition

A first-order logic (FO) sentence φ is preserved under induced substructures if for any structures $\mathcal{A}$ and $\mathcal{B}$ where $\mathcal{A}$ is an induced substructure of $\mathcal{B}$

$$\mathcal{B} \models \varphi \quad \text{implies} \quad \mathcal{A} \models \varphi$$

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 substructure.

Definition

An FO-sentence φ is universal if

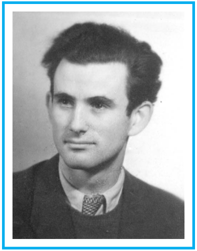
$$\varphi = \forall x_1 \cdots \forall x_k \psi,$$

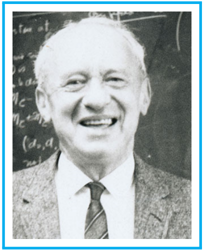
where ψ is quantifier-free.

Theorem (trivial)

Any universal sentence is preserved under induced substructures.

The Łoś-Tarski Theorem





Theorem

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

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

That is, for every structure $\mathcal{A}$, $\mathcal{A}$ satisfies φ if and only if $\mathcal{A}$ satisfies ψ. Here, $\mathcal{A}$ can be finite or infinite.

Theorem (Tait, 1959)

There is an FO*-sentence* φ *which is preserved under induced substructures in finite, i.e., for any* *finite* *structures* $\mathcal{A}$ *and* $\mathcal{B}$ *where* $\mathcal{A}$ *is an induced substructure of* $\mathcal{B}$

$$\mathcal{B} \models \varphi \quad \Longrightarrow \quad \mathcal{A} \models \varphi,$$

such that φ *is* *not* *equivalent to any universal sentence.*

Let

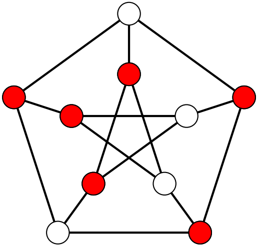
$$\varphi_k := \exists x_1 \cdots \exists x_k \forall u \forall v \left(Euv \to \bigvee_{1 \leqslant i \leqslant k} u = x_i \vee v = x_i \right).$$

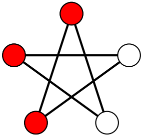
Then for any graph G

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

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 usually finite graphs.

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

Definition

Let $k \geq 1$ and $G = (V, E)$ be a graph. Then a vertex cover is a vertex subset $S \subseteq V$ such that for every $uv \in E$

$$u \in S \quad \text{or} \quad v \in S.$$

For any graph G, finite or infinite,

$$G \text{ has a vertex cover of size at most } k$$

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

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

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

1. We can write any universal sentence in CNF as

$$\psi = \forall x_1 \cdots \forall x_m \bigwedge_{i \in I} \bigvee_{j \in J_i} \delta_{ij},$$

where each δ_{ij} is an atom or a negated atom. Recall an atom is of the form

$$Ex_i x_j \quad \text{or} \quad x_i = x_j$$

for some $1 \leqslant i \leqslant j \leqslant m$.

2. ψ is equivalent to

$$\bigwedge_{i \in I} \forall x_1 \cdots \forall x_m \bigvee_{j \in J_i} \delta_{ij},$$

3. ψ is equivalent to

$$\bigwedge_{i \in I} \neg \exists x_1 \cdots \exists x_m \bigwedge_{j \in J_i} \neg \delta_{ij},$$

Lemma (folklore)

Assume

$$\chi = \exists x_1 \cdots \exists x_m \bigwedge_{j \in J} \gamma_j,$$

where every γ_j is an atom or a negated atom. Then there are graphs $H_1, \ldots, H_\ell$ such that for any graph G

$$G \models \chi \quad \iff \quad \text{one of } H_i\text{'s is an induced subgraph of } G$$

Proof.

Let $H_1, \ldots, H_\ell$ be an enumeration of all graphs H such that

- H has at most m vertices,
- and $H \models \chi$.

□

Lemma

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

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

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 $H_{i1}, \dots, H_{i\ell_i}$ such that for any graph G
$$G \models \neg \exists x_1 \cdots \exists x_m \bigwedge_{j \in J_i} \gamma_{ij} \quad \iff \quad \text{no } H_{ij} \text{ is an induced subgraph of } G.$$

$\square$

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

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

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

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

3. Let $H_{i_1}, \ldots, H_{i_{m_k}}$ be those minimal H_i's with respect to the subgraph ordering. Then for any graph G

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

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 Łoś-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 by forbidden induced subgraphs.

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

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

Lemma

Let $k \geq 1$. Then the tree-depth of a path of length k

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

Lemma

If $\mathrm{td}(G) \leqslant k$, then every path in G has length at most $2^{k-1} - 1$.

Remark

1. *The tree-width of P_k is 1.*
2. *Graphs of small tree-depth admit tree-decompositions of small width and small depth.*

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

$$\mathrm{td}(G) \leqslant k \quad \iff \quad \text{no } H_i \text{ is a subgraph of } G.$$

Remark

1. *The original proof is purely combinatorial, using Higman's Lemma on well quasi-ordering. It is non-constructive.*
2. *The result is originally about finite graphs, and the combinatorial proof does not apply to infinite graphs.*

Theorem (*C.* and Flum, 2020)

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

1. K *is definable in* FO,
2. *K is closed under subgraphs, i.e., for every $G \in \mathsf{K}$ and H a subgraph of G we have $H \in \mathsf{K}$,*

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

$$G \in \mathsf{K} \quad \Longleftrightarrow \quad \textit{no } H_i \textit{ is a subgraph of } G.$$

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 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

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

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

How to define in FO

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

For every $k \geq 1$ we define inductively φ_k expressing $\mathrm{td}(G) \leqslant k$ for a connected G.

(i) $\varphi_1 := \forall x \forall y \; x = y$.

(ii) $\varphi_{k+1} := \exists x$ "every connected component $G \setminus x$ satisfies φ_k".

Recall:

Lemma

If $\mathrm{td}(G) \leqslant k$*, then every path in* G *has length at most* $2^{k-1} - 1$*. This holds for infinite* G *as well.*

Then u and v are connected in $G \setminus x$ if and only if the following FO-sentence is true in G

$$\begin{aligned}\exists x_1 \cdots \exists x_{2^{k-1}} \Big(& \bigwedge_{1 \leqslant i \leqslant 2^{k-1}} x_i \neq x \\ & \wedge x_1 = u \wedge x_{2^{k-1}} = v) \\ & \wedge \bigwedge_{1 \leqslant i < 2^{k-1}} \big(x_i = x_{i+1} \vee E x_i x_{i+1}\big)\Big)\end{aligned}$$

Lemma

Let G and H be two graphs, finite or infinite, such that

1. $\mathrm{td}(G) \leqslant k$,
2. *H is a subgraph of G.*

Then $\mathrm{td}(H) \leqslant k$.

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 subgraphs.
4. Apply our meta-theorem to show that K can be characterized by a set of forbidden subgraphs
$$H_1, \ldots, H_m.$$

Question

How to compute $H_1, \ldots, H_m$?

Recall:

Theorem (*C.* and Flum, 2020)

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

1. K *is definable in* FO, *there is an* FO-*sentence* φ

$$\mathsf{K} = \mathrm{Mod}(\varphi) = \{G \mid \textit{graph } G \models \varphi\}.$$

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

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

Moreover, there is an algorithm which computes $H_1, \ldots, H_m$ *from* φ.

Computing the forbidden induced subgraphs from φ

1. By the Łoś-Tarski theorem there is a universal sentence ψ with
$$\models \varphi \leftrightarrow \psi.$$

2. By the completeness theorem
$$\vdash \varphi \leftrightarrow \psi,$$
that is, there is a finite proof π with
$$\pi \vdash \varphi \leftrightarrow \psi.$$

3. We enumerate all possible proofs π and universal ψ until we find $\pi \vdash \varphi \leftrightarrow \psi$.

4. We can extract a set of forbidden induced subgraphs from ψ using its normal form.

It only provides a very generic algorithm to search for the forbidden induced subgraphs, without any explicit construction and without any bound on the running time.

The explicit construction of the forbidden induced subgraphs is only known for tree-depth at most 3 [Dvorak, Giannopoulou, and Thilikos, 2012].

Why explicit construction might be hard?

The Failure of the Łoś-Tarski Theorem in Finite

Theorem (Tait, 1959)

There is an FO*-sentence which is preserved under induced substructures in finite, i.e., for any finite structures $\mathcal{A}$ and $\mathcal{B}$ where $\mathcal{A}$ is an induced substructure of $\mathcal{B}$*

$$\mathcal{B} \models \varphi \implies \mathcal{A} \models \varphi,$$

such that φ is not equivalent to any universal sentence.

Remark

Tait's examples might be viewed as colored directed graphs.

Theorem (*C.* and Flum, 2021)

There is a class K *of finite graphs satisfying the following properties.*

1. K *is closed under induced subgraphs.*
2. *There is an* FO-*sentence* φ *such that for every finite graph* G

$$G \in \mathsf{K} \iff G \models \varphi.$$

3. K *cannot be characterized by a finite set of forbidden induced subgraphs.*

1. The logic machinery cannot be applied directly to classes of finite graphs.
2. The techniques we've developed enable us to prove a number of results to explain why finding forbidden induced subgraphs might be a hard problem.

Built on [Gurevich, 1984]

Theorem ($C.$ and Flum, 2021)

For any computable function $f : \mathbb{N} \to \mathbb{N}$, *e.g.,* $f(x) = 2^{2^x}$, *there is a class* K *of graphs and an* FO*-sentence* φ *such that:*

1. $\mathrm{K} = \mathrm{Mod}(\varphi)$.
2. K *is closed under induced subgraphs.*
3. *For any forbidden induced subgraph characterization of* K *by*

   $$H_1, \dots, H_m$$

   we have

   $$\max_{i \in [m]} |H_i| \geq f(|\varphi|).$$

K has a very succinct description by FO*, but gigantic (minimal) forbidden induced subgraphs.*

For every φ we define

$$\mathsf{Mod}_{\mathsf{fin}}(\varphi) = \{ G \mid \text{finite graph } G \models \varphi \}.$$

Theorem (*C.* and Flum, 2021)

There is no algorithm that for any φ whose $\mathsf{Mod}_{\mathsf{fin}}(\varphi)$ can be characterized by a finite set of forbidden induced subgraphs computes such a set of forbidden induced subgraphs.

Compared to:

Theorem (*C.* and Flum, 2021)

There is an algorithm that for any φ whose

$$\mathsf{Mod}(\varphi) = \{ G \mid \textit{graph } G \models \varphi \}.$$

can be characterized by a finite set of forbidden induced subgraphs computes such a set of forbidden induced subgraphs.

1. We transfer Tait's Theorem and Gurevich's Theorem on arbitrary structures to graphs. In logic, this is done by FO-interpretations.
2. An FO-interpretation I translates any graph G to a structure $\mathcal{A} = \mathcal{A}_I(G)$ such that for every FO-sentence φ there is an FO-sentence φ^I with

   $$\mathcal{A} \models \varphi \iff G \models \varphi^I.$$

   Any property of $\mathcal{A}$ can be captured by a property in G.
3. We need that if G is an induced subgraph of H then $\mathcal{A}_I(G)$ is a induced substructure of $\mathcal{A}_I(H)$. This is not true for the existing FO-interpretations.
4. We introduce the notion of strongly existential interpretations, which preserves the closure of induced substructures/graphs. It requires some technical work and graph gadgets to design desired strongly existential interpretations.

(i) Forbidden induced subgraphs characterization of a class K of graphs is equivalent to
- K is closed under induced subgraphs,
- and K is definable by an FO-sentence φ (φ is not necessarily universal).

Another more complicated example is the class of graphs of bounded shrub-depth [Chen and Flum, 2020].

(ii) One can compute a set of forbidden induced subgraphs $H_1, \ldots, H_m$ for K from φ.

(iii) An important caveat is that (i) only holds when graphs can be finite or infinite. Otherwise, we've exhibited a class of finite graphs which is closed under induced subgraphs and definable in FO, but has no finite set of forbidden induced subgraphs.

(iv) For (ii) we've proved that $H_1, \ldots, H_m$ can be arbitrarily complex compared to φ. Moreover, if we only consider finite graphs, $H_1, \ldots, H_m$ cannot even be computed from φ.

1. Yijia Chen and Jörg Flum. FO-definability of shrub-depth. Conference version appeared in Proceedings of the 28th EACSL Annual Conference on Computer Science Logic (CSL'20), 15:1-15:16, 2020.

2. Yijia Chen and Jörg Flum. Forbidden induced subgraphs and the Łoś-Tarski theorem. Conference version to appear in Proceedings of the 36th Annual ACMIEEE Symposium on Logic in Computer Science (LICS'21), 2021.

Thank You!