

Sparsification Upper and Lower Bounds for Graph Problems and Not-All-Equal SAT

Bart M.P. Jansen and Astrid Pieterse



Technische Universiteit
Eindhoven
University of Technology

October 2, 2015

Kernelization

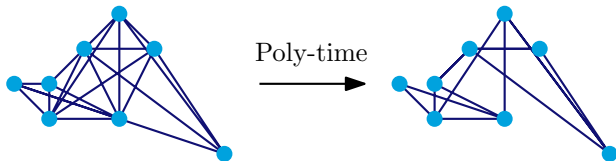
In polynomial time reduce the size of an instance

- ▶ Size depends only on the parameter

Sparsification

In **polynomial time**, create an instance that is less dense

- ▶ Has sub-quadratic number of edges
- ▶ Use kernelization framework
- ▶ Parameter: $n = |V|$



Graph problems: trivial kernel of size $O(n^2)$. Can we do better?
Use **generalized kernel** for lower bounds

Unless $NP \subseteq coNP/poly$:

- ▶ VERTEX COVER, no generalized kernel of size $O(n^{2-\epsilon})$
d-CNF-SAT, no generalized kernel of size $O(n^{d-\epsilon})$
 - Dell, van Melkebeek (J ACM14)
- ▶ TREEWIDTH, no generalized kernel of size $O(n^{2-\epsilon})$
 - Jansen (IPEC13)

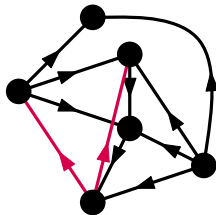
Do other graph and logic problems allow polynomial time sparsification?

Our results

4/25

Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET

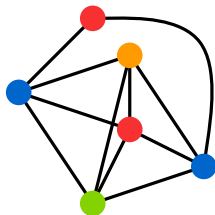


Our results

4/25

Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$

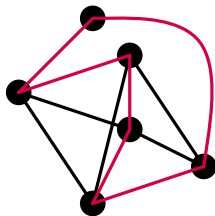


Our results

4/25

Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$
- ▶ HAMILTONIAN CYCLE

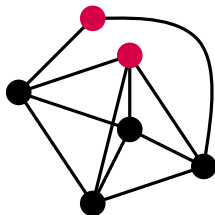


Our results

4/25

Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$
- ▶ HAMILTONIAN CYCLE
- ▶ DOMINATING SET

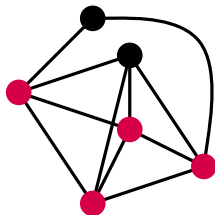


Our results

4/25

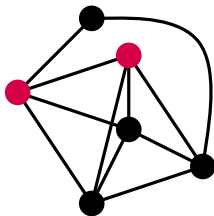
Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$
- ▶ HAMILTONIAN CYCLE
- ▶ DOMINATING SET
- ▶ NON BLOCKER
 - Kernel with $O(k)$ vertices by Dehne et al. (SOFSEM06)



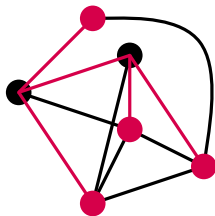
Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$
- ▶ HAMILTONIAN CYCLE
- ▶ DOMINATING SET
- ▶ NON BLOCKER
 - Kernel with $O(k)$ vertices by Dehne et al. (SOFSEM06)
- ▶ CONNECTED DOMINATING SET



Problems without a generalized kernel of size $O(n^{2-\varepsilon})$ for any $\varepsilon > 0$, unless $\text{NP} \subseteq \text{coNP/poly}$.

- ▶ FEEDBACK ARC SET
- ▶ 4-COLORING
 - And thereby k -COLORING for $k \geq 4$
- ▶ HAMILTONIAN CYCLE
- ▶ DOMINATING SET
- ▶ NON BLOCKER
 - Kernel with $O(k)$ vertices by Dehne et al. (SOFSEM06)
- ▶ CONNECTED DOMINATING SET
- ▶ MAXIMUM LEAF SPANNING TREE
 - Kernel with $O(k)$ vertices by Estivill-Castro et al. (ACiD05)



(Non)-sparsifiability of CNF-formulas

- ▶ d -NAE-SAT has a generalized kernel of size $O(n^{d-1})$
 - Matches known lower bound

(Non)-sparsifiability of CNF-formulas

- ▶ d -NAE-SAT has a generalized kernel of size $O(n^{d-1})$
 - Matches known lower bound

Compare to d -CNF-SAT, $O(n^{d-\varepsilon})$ kernel unlikely

d-NAE-SAT

- ▶ Input: CNF-formula \mathcal{F} , each clause contains at most d literals.

$$\mathcal{F} = \underbrace{(x \vee \neg y \vee \dots \vee z)}_{\text{Clause, } \leq d \text{ literals}} \wedge (\neg x \vee \neg z \vee \dots \vee \neg y) \wedge \dots$$

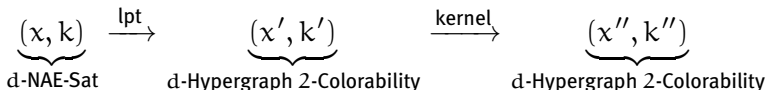
- ▶ Parameter: The number of variables n .
- ▶ Question: Find a truth assignment such that every clause contains a **true** and a **false** literal?

Not-All-Equal Satisfiability: kernel bounds

7/25

No generalized kernel of size $O(n^{d-1-\epsilon})$, unless $NP \subseteq coNP/poly$.

- ▶ Linear parameter transformation d -CNF-SAT to $(d+1)$ -NAE-SAT
 - Shown by Jansen et al. (InformComput13)
- ▶ Is this a tight lower bound?
 - Not trivial.
- ▶ Provide a generalized kernel via d -HYPERGRAPH 2-COLORABILITY



d-HYPERGRAPH 2-COLORABILITY

- ▶ Input: Hypergraph, every edge contains **at most** d vertices.
- ▶ Parameter: The number of vertices n .
- ▶ Question: Color each vertex with red/blue such that every edge contains a red and a blue vertex?

Kernel

d-HYPERGRAPH 2-COLORABILITY has a kernel with $O(n^{d-1})$ edges.

Construction

- ▶ Let every edge have **exactly** d vertices
- ▶ Edges e_1, \dots, e_m
- ▶ Enumerate all size $d - 1$ subsets of V as S_1, S_2, \dots, S_ℓ
- ▶ Create $(0, 1)$ -matrix M

$$\begin{array}{c} S_1 \\ S_2 \\ \dots \\ S_\ell \end{array} \begin{pmatrix} e_1 & e_2 & \dots & e_m \\ S_1 \subseteq e_1 & S_1 \subseteq e_2 & \dots & S_1 \subseteq e_m \\ S_2 \subseteq e_1 & S_2 \subseteq e_2 & \dots & S_2 \subseteq e_m \\ \dots & \dots & \dots & \dots \\ S_\ell \subseteq e_1 & S_\ell \subseteq e_2 & \dots & S_\ell \subseteq e_m \end{pmatrix}$$

- ▶ Compute a **base** of the columns of this matrix.
- ▶ This results in a subset of the edges of G

Size

- ▶ Matrix M has at most $\binom{n}{d-1} \leq n^{d-1}$ rows
- ▶ Any base of M contains at most n^{d-1} edges

Correctness

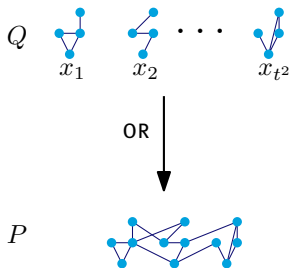
- ▶ If G is 2-colorable, so is the reduced graph
- ▶ Edges in the base are split implies removed edges split
 - Proof uses a lemma by Lovász (1976)
 - **Critical 3-chromatic hypergraphs** have at most $O(n^{d-1})$ edges
 - Transferred proof idea into kernel

Degree-2 cross-composition to P gives kernel lower bound $O(n^2)$

Degree-2 cross-composition

By Bodlaender et al. (SIDMA14)

- ▶ Start from any NP-hard problem Q
- ▶ Give a polynomial time algorithm
- ▶ Input: t^2 **similar** instances of Q
- ▶ Output: An instance (y, k) of P , where
 - $k = O(t \cdot \max |x_i|^c)$
 - (y, k) is a logical OR of the inputs



Goal

Prove that there is no kernel of size $O(n^{2-\varepsilon})$ for 4-COLORING.

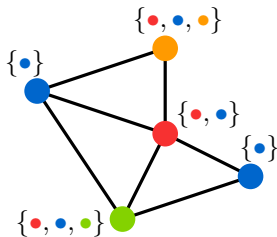
Goal

Prove that there is no kernel of size $O(n^{2-\varepsilon})$ for 4-COLORING.

4-LIST COLORING

We use list coloring with 4 colors in total, instead of 4-COLORING.

- ▶ Vertex has list of allowed colors
 - Subset of $\{\text{red}, \text{green}, \text{blue}, \text{orange}\}$



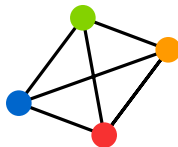
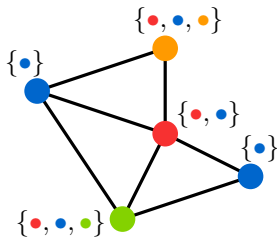
Goal

Prove that there is no kernel of size $O(n^{2-\epsilon})$ for 4-COLORING.

4-LIST COLORING

We use list coloring with 4 colors in total, instead of 4-COLORING.

- ▶ Vertex has list of allowed colors
 - Subset of $\{\text{red}, \text{green}, \text{blue}, \text{orange}\}$
- ▶ Transform back to 4-COLORING



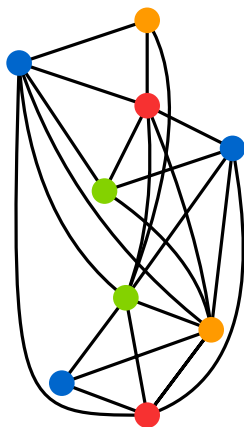
Goal

Prove that there is no kernel of size $O(n^{2-\epsilon})$ for 4-COLORING.

4-LIST COLORING

We use list coloring with 4 colors in total, instead of 4-COLORING.

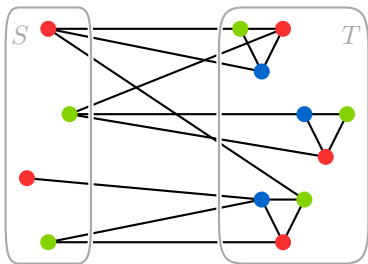
- ▶ Vertex has list of allowed colors
 - Subset of $\{\text{red}, \text{green}, \text{blue}, \text{orange}\}$
- ▶ Transform back to 4-COLORING



NP-hard starting problem

2-3-COLORING ON TRIANGLE-SPLIT GRAPHS

- ▶ Input: Graph $G = (S \cup T, E)$ where S is an independent set and T consists of disjoint triangles.
- ▶ Question: Does G have a proper 3-coloring, such that S is colored using only 2 colors?
 - We call this a *2-3-coloring* of G .



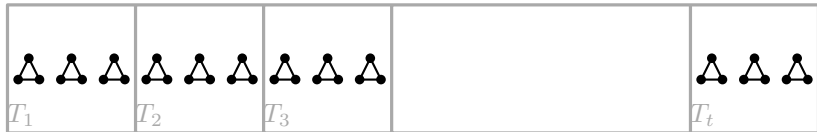
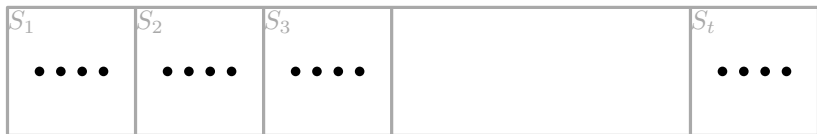
4-Coloring: Cross-composition

14/25

- ▶ Assume we have t^2 instances
- ▶ Let $|S| = n$ and $|T| = 3m$ for all instances
- ▶ Construct a cross-composition
 - At most $O(t \cdot (n + m))$ vertices
- ▶ We cannot copy all vertices
 - But we can keep all edges
 - Method introduced by Dell and Marx (SODA12)
- ▶ Enumerate instances as X_{ij} , where $i = 1, \dots, t$

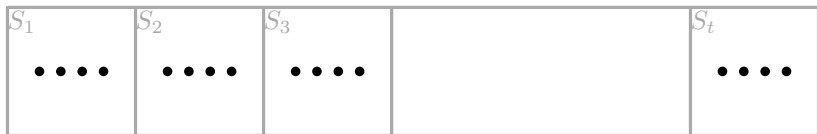
4-Coloring: First idea

15/25

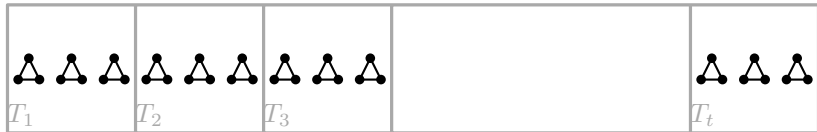


4-Coloring: First idea

15/25

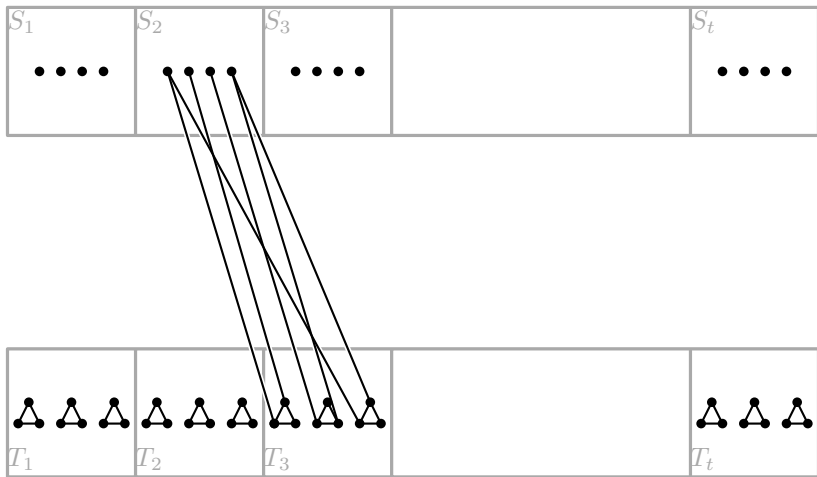


Instance X_{23} ?



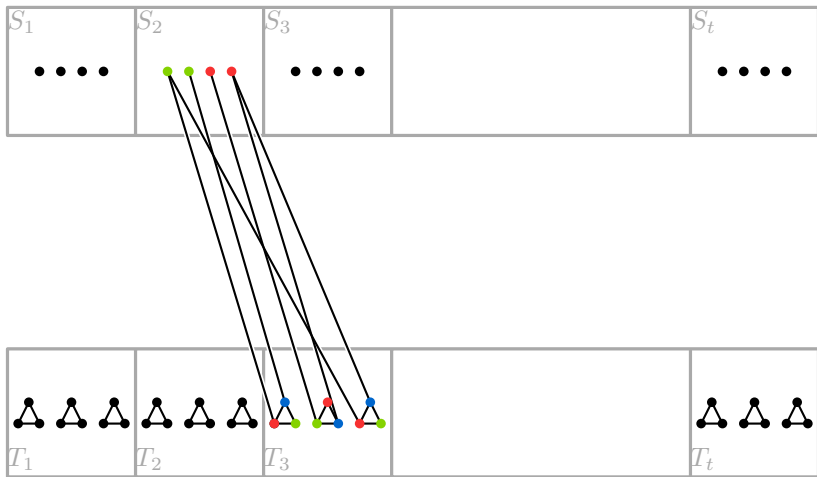
4-Coloring: First idea

15/25



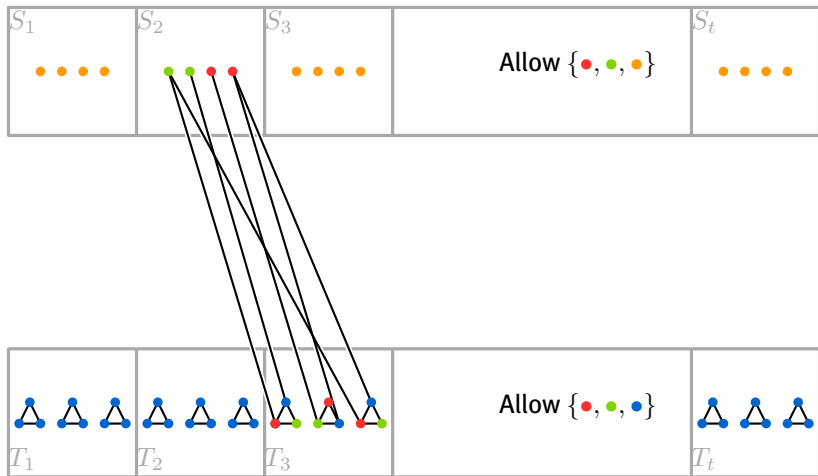
4-Coloring: First idea

15/25



4-Coloring: First idea

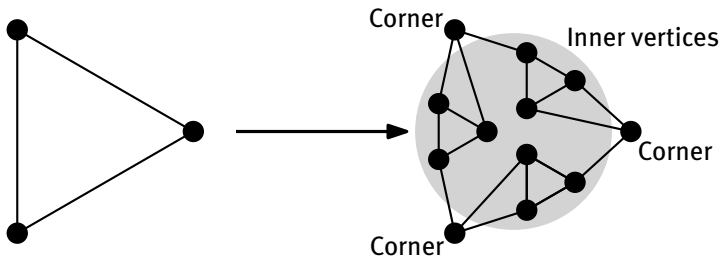
15/25



4-Coloring: Triangular gadget

16/25

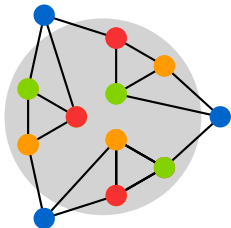
This is not a valid coloring of the triangles. Replace them:



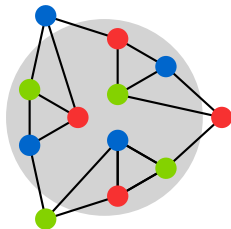
Allow the new vertices to be red, green, orange, or blue.

Useful properties

We can color all corners blue

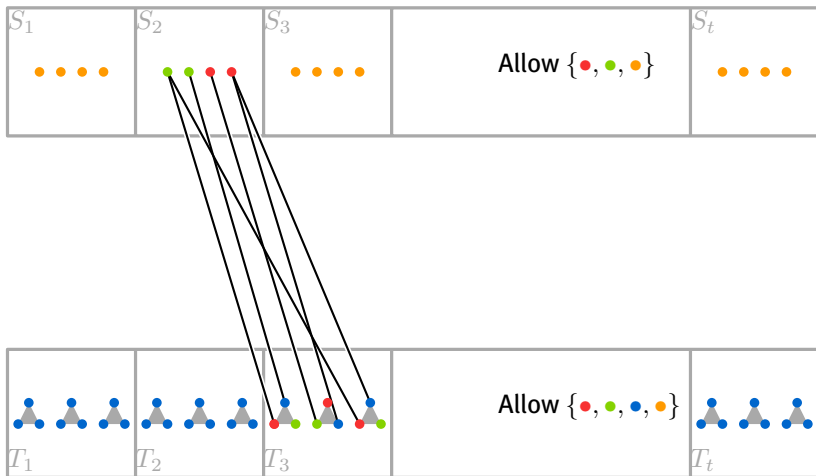


If 3-colored: ordinary triangle



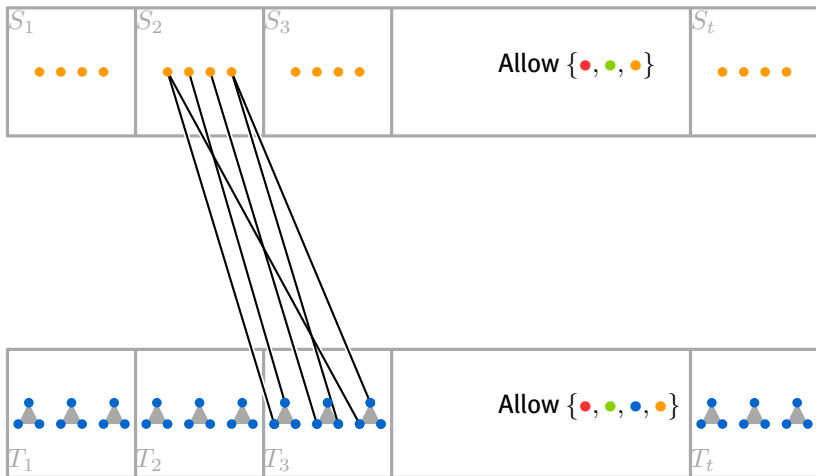
4-Coloring: Improved

18/25



4-Coloring: Improved

18/25



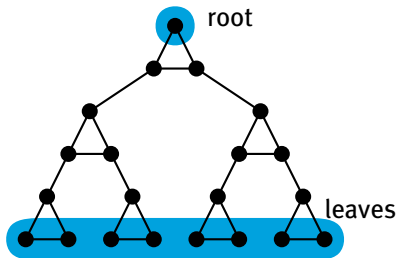
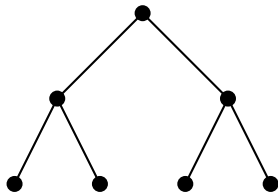
4-Coloring: Treegadgets

19/25

Ensure one S_i and one T_j are colored without orange

Treegadget

Complete binary tree where every vertex is replaced by a triangle:



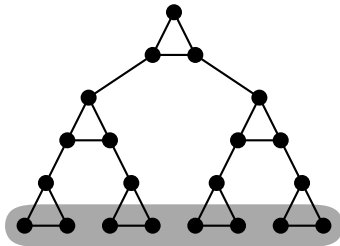
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



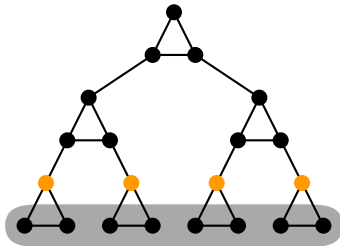
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



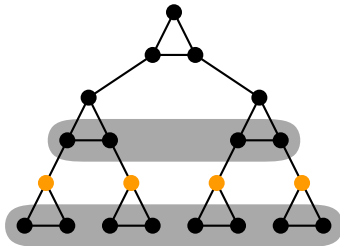
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



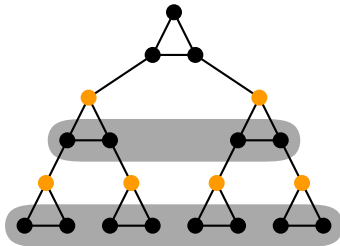
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



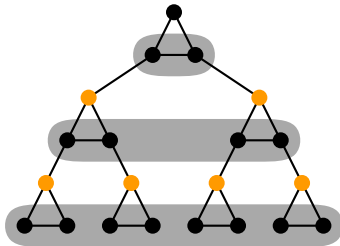
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



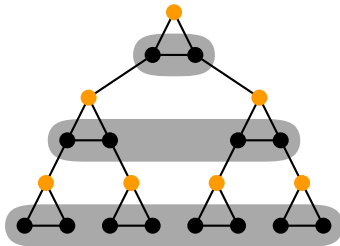
4-Coloring: Treegadgets

20/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 1

If none of the leaves is colored orange, the root must be orange.



4-Coloring: Treegadgets

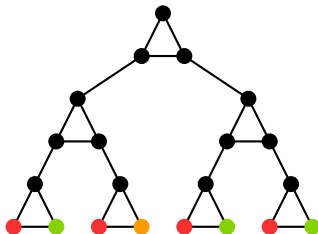
21/25

Suppose we want to 3-color a gadget, using red, green and orange.

Property 2

We can extend a coloring of the leaves, to color the entire tree.

- ▶ If at least one of the leaves is orange, the root can be red or green.



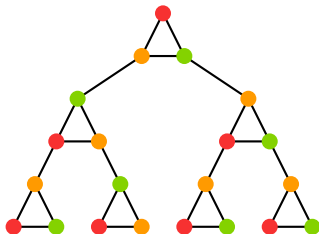
4-Coloring: Treegadgets

Suppose we want to 3-color a gadget, using red, green and orange.

Property 2

We can extend a coloring of the leaves, to color the entire tree.

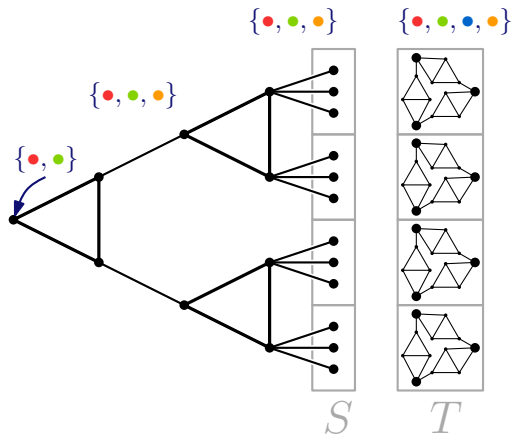
- ▶ If at least one of the leaves is orange, the root can be red or green.



4-Coloring: Cross-composition

22/25

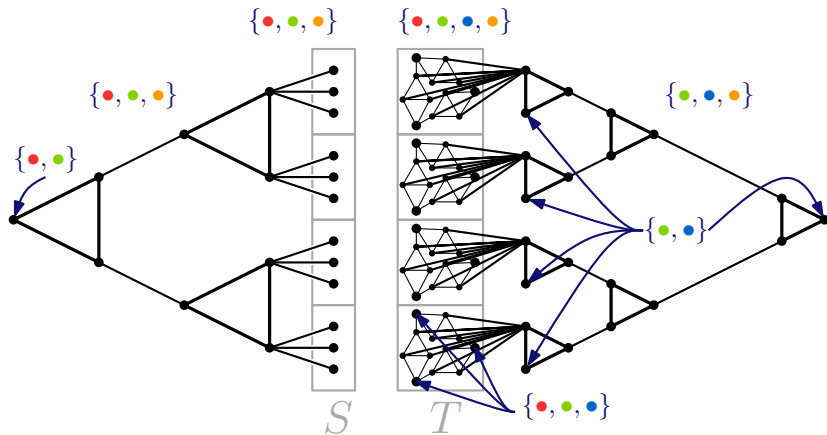
Ensure one group in S is colored with red and green.



4-Coloring: Cross-composition

23/25

Ensure one group in T is colored with red, green and orange.



- ✓ The number of vertices of G is allowed: $O(t \cdot \max |X_{ij}|)$.
- ✓ Can be done in polynomial time.
- ✓ If some X_{ij} is 2-3-colorable, G is 4-colorable.
- ✓ If G is 4-colorable, there exists an X_{ij} that is 2-3-colorable.

Theorem

4-COLORING does not have a generalized kernel of size $O(n^{2-\epsilon})$ for $\epsilon > 0$, unless $NP \subseteq coNP/poly$.

- ▶ Non-sparsifiability of several classic graph problems
 - Implications for k -NON BLOCKER and k -MAX LEAF SPANNING TREE
- ▶ Generalized kernel for NAE-SAT
 - One of the first non-trivial sparsifications.

Open problems

- ▶ Does 3-COLORING allow sparsification?
- ▶ Sparsification of NP-hard problem on general graphs
- ▶ Sparsification of edge-based problems
 - Max Cut

- ▶ Non-sparsifiability of several classic graph problems
 - Implications for k -NON BLOCKER and k -MAX LEAF SPANNING TREE
- ▶ Generalized kernel for NAE-SAT
 - One of the first non-trivial sparsifications.

Open problems

- ▶ Does 3-COLORING allow sparsification?
- ▶ Sparsification of NP-hard problem on general graphs
- ▶ Sparsification of edge-based problems
 - Max Cut

Thank you for your attention