Fixed-Parameter Linear-Time Algorithms for Graph and Hypergraph Problems Arising in Industrial Applications

René van Bevern

Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany

Wissenschaftliche Aussprache, 17. Juni 2014

Fixed-parameter linear-time algorithms

NP-hard problems considered in the thesis

d-Hitting Set

 Race condition detection in parallel Java programs.
[O'Callahan and Choi, ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming 2003]

Dag Partitioning

 Real-time tracking of trends and topics on the internet. [Leskovec, Backstrom, and Kleinberg, ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2009]
[Suen, Huang, Eksombatchai, Sosic, and Leskovec, International Conference on World Wide Web 2013]

2-Union Independent Set

 Job scheduling, e.g. in steel manufacturing. [Höhn, König, Möhring, and Lübbecke, Management Science 57, 2011]

Hypergraph Cutwidth

 Automatic testing of circuits. [Prasad, Chong, and Keutzer, Design Automation Conference 1999]

Fixed-parameter linear-time algorithms

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain parameters of the input are small.

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain **parameters** of the input are **small**.
- **Approach:** Solve problems in $f(k) \cdot poly(n)$ time for some parameter $k \rightsquigarrow$ **fixed-parameter algorithms**.

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain parameters of the input are small.
- **Approach:** Solve problems in $f(k) \cdot poly(n)$ time for some parameter $k \rightsquigarrow$ **fixed-parameter algorithms**.

Common lines of research:

- Make f(k) smaller, e.g. improve it from k^k to 2^k .
- ► Find fixed-parameter algorithms for smaller parameters.

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain parameters of the input are small.
- **Approach:** Solve problems in $f(k) \cdot poly(n)$ time for some parameter $k \rightsquigarrow$ **fixed-parameter algorithms**.

Common lines of research:

- Make f(k) smaller, e.g. improve it from k^k to 2^k .
- Find fixed-parameter algorithms for smaller parameters. **Still:** Algorithms running in $O(2^k \cdot n^6)$ time.

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain parameters of the input are small.
- **Approach:** Solve problems in $f(k) \cdot poly(n)$ time for some parameter $k \rightsquigarrow$ **fixed-parameter algorithms**.

Common lines of research:

- Make f(k) smaller, e.g. improve it from k^k to 2^k .
- Find fixed-parameter algorithms for smaller parameters. **Still:** Algorithms running in $O(2^k \cdot n^6)$ time.

Focus change: Solve problems in **linear time** for constant parameter values \rightsquigarrow **fixed-parameter linear-time algorithms**.

- Challenge: No polynomial-time algorithms for NP-hard problems known.
 - **Goal:** Solve NP-hard problems efficiently if certain parameters of the input are small.
- **Approach:** Solve problems in $f(k) \cdot poly(n)$ time for some parameter $k \rightsquigarrow$ **fixed-parameter algorithms**.

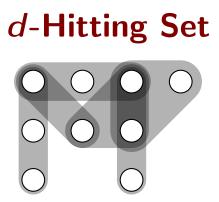
Common lines of research:

- Make f(k) smaller, e.g. improve it from k^k to 2^k .
- Find fixed-parameter algorithms for smaller parameters. **Still:** Algorithms running in $O(2^k \cdot n^6)$ time.

Focus change: Solve problems in **linear time** for constant parameter values \rightsquigarrow **fixed-parameter linear-time algorithms**.

Back to the roots: [Bodlaender, SIAM Journal on Computing 25, 1996: "A Linear-Time Algorithm for Finding Tree-Decompositions of Small Treewidth"]

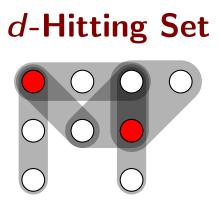
Fixed-parameter linear-time algorithms



Input: A hypergraph H = (V, E) with a set V of vertices, a set $E \subseteq 2^V$ of hyperedges, each of cardinality at most a constant d, and a natural number k.

Question: Is there a hitting set $S \subseteq V$ of size at most k, that is, $\forall e \in E : S \cap E \neq \emptyset$?

Fixed-parameter linear-time algorithms



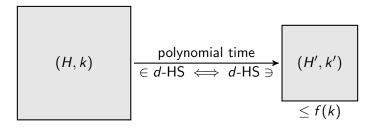
Input: A hypergraph H = (V, E) with a set V of vertices, a set $E \subseteq 2^V$ of hyperedges, each of cardinality at most a constant d, and a natural number k.

Question: Is there a hitting set $S \subseteq V$ of size at most k, that is, $\forall e \in E : S \cap E \neq \emptyset$?

Fixed-parameter linear-time algorithms

Problem kernelization

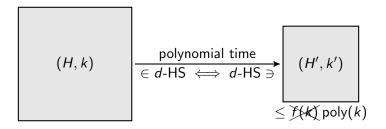
d-HS is NP-hard \rightsquigarrow data reduction, **problem kernelization**:



Fixed-parameter linear-time algorithms

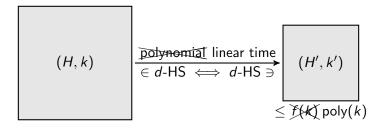
Problem kernelization

d-HS is NP-hard \rightsquigarrow data reduction, **problem kernelization**:



Problem kernelization

d-HS is NP-hard \rightsquigarrow data reduction, **problem kernelization**:



No O($k^{d-\varepsilon}$)-size problem kernel.

[Dell and van Melkebeek, ACM Symposium on Theory of Computing 2010]

O(*k^d*)-size problem kernels for *d*-HS in polynomial time. [Flum and Grohe, Parameterized Complexity Theory, 2006] [Damaschke, Theoretical Computer Science 351, 2006] [S. Kratsch, Algorithmica 62, 2012]

O(k^{d-1})-vertex problem kernels for *d*-HS in polynomial time. [Abu-Khzam, Journal of Computer and System Sciences 76, 2010] [Moser, Dissertation, Friedrich-Schiller-Universität Jena, 2010]

No O($k^{d-\varepsilon}$)-size problem kernel.

[Dell and van Melkebeek, ACM Symposium on Theory of Computing 2010]

O(k^d)-size problem kernels for d-HS in polynomial time. [Flum and Grohe, Parameterized Complexity Theory, 2006] [Damaschke, Theoretical Computer Science 351, 2006] [S. Kratsch, Algorithmica 62, 2012]

O(k^{d-1})-vertex problem kernels for *d*-HS in polynomial time. [Abu-Khzam, Journal of Computer and System Sciences 76, 2010] [Moser, Dissertation, Friedrich-Schiller-Universität Jena, 2010]

New: $O(k^d)$ -size problem kernel in O(|V| + |E|) time.

Fixed-parameter linear-time algorithms

No O($k^{d-\varepsilon}$)-size problem kernel.

[Dell and van Melkebeek, ACM Symposium on Theory of Computing 2010]

O(k^d)-size problem kernels for d-HS in polynomial time. [Flum and Grohe, Parameterized Complexity Theory, 2006] [Damaschke, Theoretical Computer Science 351, 2006] [S. Kratsch, Algorithmica 62, 2012]

O(k^{d-1})-vertex problem kernels for *d*-HS in polynomial time. [Abu-Khzam, Journal of Computer and System Sciences 76, 2010] [Moser, Dissertation, Friedrich-Schiller-Universität Jena, 2010]

New: $O(k^d)$ -size problem kernel in O(|V| + |E|) time.

Running time and size are essentially optimal.

No O($k^{d-\varepsilon}$)-size problem kernel.

[Dell and van Melkebeek, ACM Symposium on Theory of Computing 2010]

O(k^d)-size problem kernels for d-HS in polynomial time. [Flum and Grohe, Parameterized Complexity Theory, 2006] [Damaschke, Theoretical Computer Science 351, 2006] [S. Kratsch, Algorithmica 62, 2012]

O(k^{d-1})-vertex problem kernels for *d*-HS in polynomial time. [Abu-Khzam, Journal of Computer and System Sciences 76, 2010] [Moser, Dissertation, Friedrich-Schiller-Universität Jena, 2010]

New: $O(k^d)$ -size problem kernel in O(|V| + |E|) time.

- Running time and size are essentially optimal.
- *d*-HS solvable in $O(d^k + |V| + |E|)$ time.

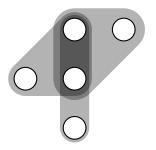
Sunflowers

A family P of sets that pairwise only intersect in a **core** C is a **sunflower**. The sets in P are **petals**.



Sunflowers

A family P of sets that pairwise only intersect in a **core** C is a **sunflower**. The sets in P are **petals**.



Theorem: If a *d*-uniform hypergraph *H* has more than $k^d \cdot d!$ hyperedges, then it contains a sunflower of size *k* that can be found in polynomial time.

[Erdős and Rado, Journal of the London Mathematical Society 35(1), 1960]

[Flum and Grohe, Parameterized Complexity Theory, 2006]

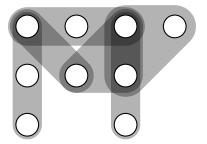
Fixed-parameter linear-time algorithms

Given a d-HS instance (H, k), repeatedly

- try to **find** a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]

Given a d-HS instance (H, k), repeatedly

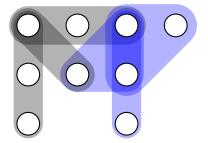
- try to find a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]



We try to kernelize this instance for k = 1.

Given a d-HS instance (H, k), repeatedly

- try to find a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]

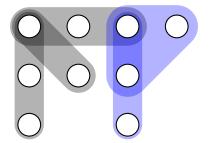


Find sunflower of size k + 2 = 3.

Fixed-parameter linear-time algorithms

Given a d-HS instance (H, k), repeatedly

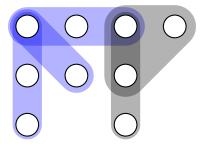
- try to **find** a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]



Remove one petal.

Given a d-HS instance (H, k), repeatedly

- try to find a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]

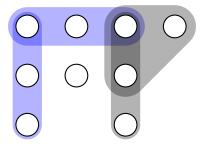


Find sunflower of size k + 2 = 3.

Fixed-parameter linear-time algorithms

Given a d-HS instance (H, k), repeatedly

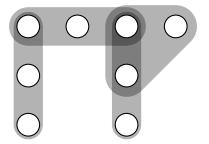
- try to find a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]



Remove one petal.

Given a d-HS instance (H, k), repeatedly

- try to find a sunflower with k + 2 petals,
- ▶ remove one of its petals from *H*. [S. Kratsch, Algorithmica 63, 2012]



Resulting problem kernel.

Fixed-parameter linear-time algorithms

Given a *d*-HS instance (H, k), create an empty hypergraph H'. Then, **for each** hyperedge *e* of *H*:

- if e does not contain the core of a (k+1)-petal sunflower in H'
 - add e to H'.

Want this in constant time.

Given a *d*-HS instance (H, k), create an empty hypergraph H'. Then, **for each** hyperedge *e* of *H*:

- ▶ if $\forall C \subseteq e$: C is not the core of a (k+1)-petal sunflower in H'
 - add e to H'.

Unfortunately NP-hard to decide.

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - ▶ **add** *e* to *H*′.

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - add e to H'.
 - ▶ for each $C \subseteq e$, if $(e \cup \text{sunflower}[C])$ is a sunflower, then
 - ▶ add *e* to sunflower[*C*]

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - add e to H'.
 - ▶ for each $C \subseteq e$, if $(e \cup \text{sunflower}[C])$ is a sunflower, then
 - ▶ add *e* to sunflower[*C*]
 - for each $v \in e \setminus C$, set used $[C][v] \leftarrow$ true

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - add e to H'.
 - for each $C \subseteq e$, if $\forall v \in e \setminus C$: used[C][v] = false, then
 - ▶ add *e* to sunflower[*C*]
 - for each $v \in e \setminus C$, set used $[C][v] \leftarrow$ true

Given a *d*-HS instance (H, k), create an empty hypergraph H'. Then, for each hyperedge *e* of *H*:

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - ▶ **add** *e* to *H*′.
 - ▶ for each $C \subseteq e$, if $\forall v \in e \setminus C$: used[C][v] = false, then
 - ▶ add *e* to sunflower[*C*]
 - for each $v \in e \setminus C$, set used $[C][v] \leftarrow$ true

sunflower[C] and used[C] can be accessed in $O(d) \subseteq O(1)$ time using a prefix tree.

Given a *d*-HS instance (H, k), create an empty hypergraph H'. Then, **for each** hyperedge *e* of *H*:

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - ▶ **add** *e* to *H*′.
 - ▶ for each $C \subseteq e$, if $\forall v \in e \setminus C$: used[C][v] = false, then
 - ▶ add *e* to sunflower[*C*]
 - for each $v \in e \setminus C$, set used $[C][v] \leftarrow$ true

sunflower[C] and used[C] can be accessed in $O(d) \subseteq O(1)$ time using a prefix tree.

Running time: $O(d|V| + 2^d d|E|)$. **Memory:** $\Theta(|V| \cdot |E|)$.

Fixed-parameter linear-time algorithms

Given a *d*-HS instance (H, k), create an empty hypergraph H'. Then, **for each** hyperedge *e* of *H*:

- if $\forall C \subseteq e$: $|\operatorname{sunflower}[C]| \leq k$, then
 - ▶ **add** *e* to *H*′.
 - ▶ for each $C \subseteq e$, if $\forall v \in e \setminus C$: used[C][v] = false, then
 - ▶ add *e* to sunflower[*C*]
 - for each $v \in e \setminus C$, set used $[C][v] \leftarrow$ true

sunflower[C] and used[C] can be accessed in $O(d) \subseteq O(1)$ time using a prefix tree.

Running time: $O(d|V| + 2^d d|E|)$. **Memory:** $\Theta(|V| \cdot |E|)$.

Prefix tree has to be initialized incompletely and carefully.

Fixed-parameter linear-time algorithms

Further results on *d*-Hitting Set

Experimental results

 Instances from a problem arising in radio frequency allocation. [Sorge, Moser, Niedermeier, and Weller, Conference on Integer Programming and Combinatorial Optimization 2012]

[Babcock, Bell Systems Technical Journal 32, 1953]

Further results on *d*-Hitting Set

Experimental results

 Instances from a problem arising in radio frequency allocation. [Sorge, Moser, Niedermeier, and Weller, Conference on Integer Programming and Combinatorial Optimization 2012]

[Babcock, Bell Systems Technical Journal 32, 1953]

▶ 4-HS with $20 \cdot 10^6$ hyperedges processed in about five minutes.

Further results on *d*-Hitting Set

Experimental results

 Instances from a problem arising in radio frequency allocation. [Sorge, Moser, Niedermeier, and Weller, Conference on Integer Programming and Combinatorial Optimization 2012]

[Babcock, Bell Systems Technical Journal 32, 1953]

- 4-HS with $20 \cdot 10^6$ hyperedges processed in about five minutes.
- Prefix trees for sunflower[] and used[] are faster, but balanced trees have linear memory usage.

Further results on *d*-Hitting Set

Experimental results

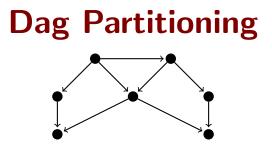
 Instances from a problem arising in radio frequency allocation. [Sorge, Moser, Niedermeier, and Weller, Conference on Integer Programming and Combinatorial Optimization 2012]

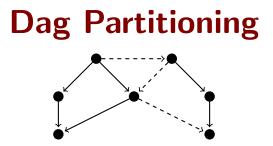
[Babcock, Bell Systems Technical Journal 32, 1953]

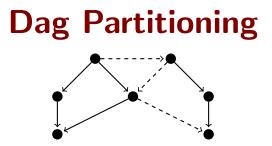
- ▶ 4-HS with $20 \cdot 10^6$ hyperedges processed in about five minutes.
- Prefix trees for sunflower[] and used[] are faster, but balanced trees have linear memory usage.

Speed up kernels of Abu-Khzam and Moser:

 ► O(k^{d-1})-vertex problem kernel in O(|V| + |E| + k^{1.5d}) time. [Abu-Khzam, Journal of Computer and System Sciences 76, 2010] [Moser, Dissertation, Friedrich-Schiller-Universität Jena, 2010]

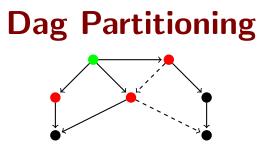






Algorithmic results:

- $O(2^k \cdot (n+m))$ time algorithm.
- Linear-time data reduction.



Algorithmic results:

- $O(2^k \cdot (n+m))$ time algorithm.
- Linear-time data reduction.

Strategy: First solve problem for all out-neighbors of v, then for v.

 $O(2^k \cdot (n+m))$ time algorithm on simulated citation networks:

 $O(2^k \cdot (n+m))$ time algorithm on simulated citation networks:

► Instances with m ≥ 10⁷ arcs and k ≤ 190 arc deletions solvable in five minutes.

 $O(2^k \cdot (n+m))$ time algorithm on simulated citation networks:

- ► Instances with m ≥ 10⁷ arcs and k ≤ 190 arc deletions solvable in five minutes.
- Without the linear-time data reduction, no instance solvable in less than an hour.

 $O(2^k \cdot (n+m))$ time algorithm on simulated citation networks:

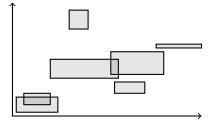
- ► Instances with m ≥ 10⁷ arcs and k ≤ 190 arc deletions solvable in five minutes.
- Without the linear-time data reduction, no instance solvable in less than an hour.

Comparison with known heuristic.

[Leskovec, Backstrom, and Kleinberg, ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2009]

- ▶ Heuristic more than a factor of 2.5 off the optimum.
- Algorithm only runs fast where the heuristic gives optimal solutions.

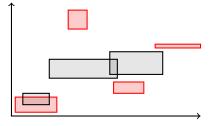
2-Union Independent Set



Task: Find **at least** k = 4 rectangles whose projections onto neither axis intersect.

Fixed-parameter linear-time algorithms

2-Union Independent Set



Task: Find **at least** k = 4 rectangles whose projections onto neither axis intersect.

Fixed-parameter linear-time algorithms



Task: Find at least k = 3 rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **one** of γ **strips**.



Task: Find at least k = 3 rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **one** of γ **strips**.

Known: Solvable in $O(2^{\gamma}\gamma \cdot n)$ time. [Halldórsson and Karlsson, International Workshop on Graph-Theoretic Concepts in Computer Science 2006]



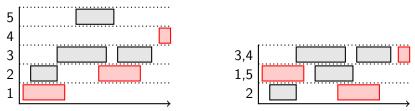
Task: Find at least k = 3 rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **one** of γ **strips**.

Known: Solvable in $O(2^{\gamma}\gamma \cdot n)$ time. [Halldórsson and Karlsson, International Workshop on Graph-Theoretic Concepts in Computer Science 2006]

New: O(5.5^k $k \cdot n$) time, where $k \leq \gamma$.

Fixed-parameter linear-time algorithms



Task: Find at least k = 3 rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **one** of γ **strips**.

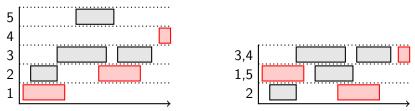
Known: Solvable in $O(2^{\gamma}\gamma \cdot n)$ time. [Halldórsson and Karlsson, International Workshop on Graph-Theoretic Concepts in Computer Science 2006]

New: O(5.5^k $k \cdot n$) time, where $k \leq \gamma$.

Trick: Randomly move strip $i \in \{1, ..., \gamma\}$ to strip $j \in \{1, ..., k\}$.

• Then use $O(2^{\gamma}\gamma \cdot n)$ -time algorithm with $\gamma = k$.

Fixed-parameter linear-time algorithms



Task: Find at least k = 3 rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **one** of γ **strips**.

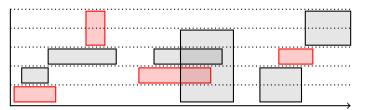
Known: Solvable in $O(2^{\gamma}\gamma \cdot n)$ time. [Halldórsson and Karlsson, International Workshop on Graph-Theoretic Concepts in Computer Science 2006]

New: O(5.5^k $k \cdot n$) time, where $k \leq \gamma$.

Trick: Randomly move strip $i \in \{1, ..., \gamma\}$ to strip $j \in \{1, ..., k\}$.

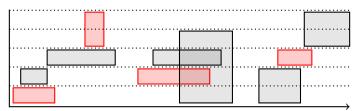
- Then use $O(2^{\gamma}\gamma \cdot n)$ -time algorithm with $\gamma = k$.
- Repeat $O(e^k \cdot |\ln \varepsilon|)$ times for error probability $\leq \varepsilon$.

Fixed-parameter linear-time algorithms



Task: Find **at least** *k* rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **any subset** of γ **strips** (we say that one axis is γ -compact).

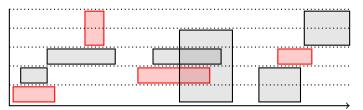


Task: Find **at least** *k* rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **any subset** of γ **strips** (we say that one axis is γ -compact).

New: $O(2^{\gamma}\gamma \cdot n)$ time algorithm.

Generalization of Halldórsson and Karlsson's algorithm.

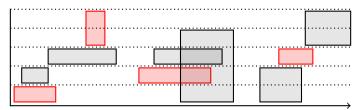


Task: Find **at least** *k* rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **any subset** of γ **strips** (we say that one axis is γ -compact).

New: $O(2^{\gamma}\gamma \cdot n)$ time algorithm.

- Generalization of Halldórsson and Karlsson's algorithm.
- γ -compact representation for minimum γ in linear time.



Task: Find **at least** *k* rectangles not intersecting on any axis.

Special case: Each rectangle vertically fills **any subset** of γ **strips** (we say that one axis is γ -compact).

New: $O(2^{\gamma}\gamma \cdot n)$ time algorithm.

- Generalization of Halldórsson and Karlsson's algorithm.
- γ -compact representation for minimum γ in linear time.

Experiments on random rectangles:

• $n \le 0.6 \cdot 10^6$ rectangles in $\gamma \le 15$ strips solved in five minutes.

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Insights from experiments:

- Simpler algorithms.
- Provable speed-ups.
- We are on the right way.

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Insights from experiments:

- Simpler algorithms.
- Provable speed-ups.
- We are on the right way.

Challenges:

Parameter race for fixed-parameter linear-time algorithms.

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Insights from experiments:

- Simpler algorithms.
- Provable speed-ups.
- We are on the right way.

Challenges:

- Parameter race for fixed-parameter linear-time algorithms.
- Linear-time transformations to problems with well-tuned solvers (SAT, ILP)

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Insights from experiments:

- Simpler algorithms.
- Provable speed-ups.
- We are on the right way.

Challenges:

- Parameter race for fixed-parameter linear-time algorithms.
- Linear-time transformations to problems with well-tuned solvers (SAT, ILP) + linear-time data reduction.

Aspects brought into fixed-parameter algorithms:

- Returning to the initial goal—efficient algorithms.
- Data structures are suddenly important.

Insights from experiments:

- Simpler algorithms.
- Provable speed-ups.
- We are on the right way.

Challenges:

- Parameter race for fixed-parameter linear-time algorithms.
- Linear-time transformations to problems with well-tuned solvers (SAT, ILP) + linear-time data reduction.
- Lower bounds.