Practise what you preach: sparsity in the real world.

Felix Reidl
felix.reidl@gmail.com

ASSG'17

Part I

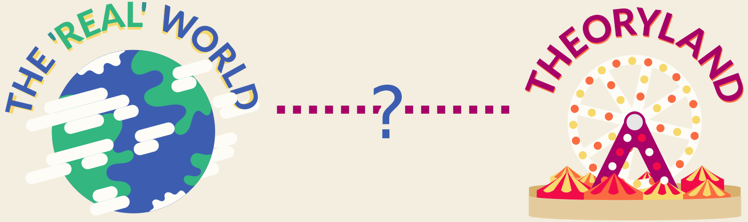
Theory vs Practise

A bit of (hard) introspection

"About ten years ago, some computer scientists came by and said they heard we have some really cool problems. They showed that the problems are NP-complete and went away!"

–Joseph Felsenstein in 1997

Downey RG, Fellows MR, Stege U. **Computational tractability: The view from mars**. Bulletin of the EATCS. 1999 Oct;69:73-97.



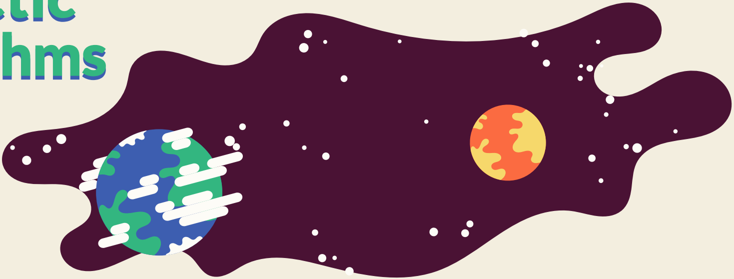
Our reputation

Small "tricks"

Our reputation

Small "tricks"

Galactic Algorithms



A communication problem

“Practical algorithm”

Polynomial time

Linear time

Short description

Logspace

Reduces to
well-known problem

Easy to understand

A communication problem

"Practical algorithm"

Polynomial time

Linear time

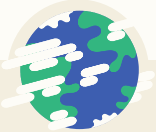
Short description

Logspace

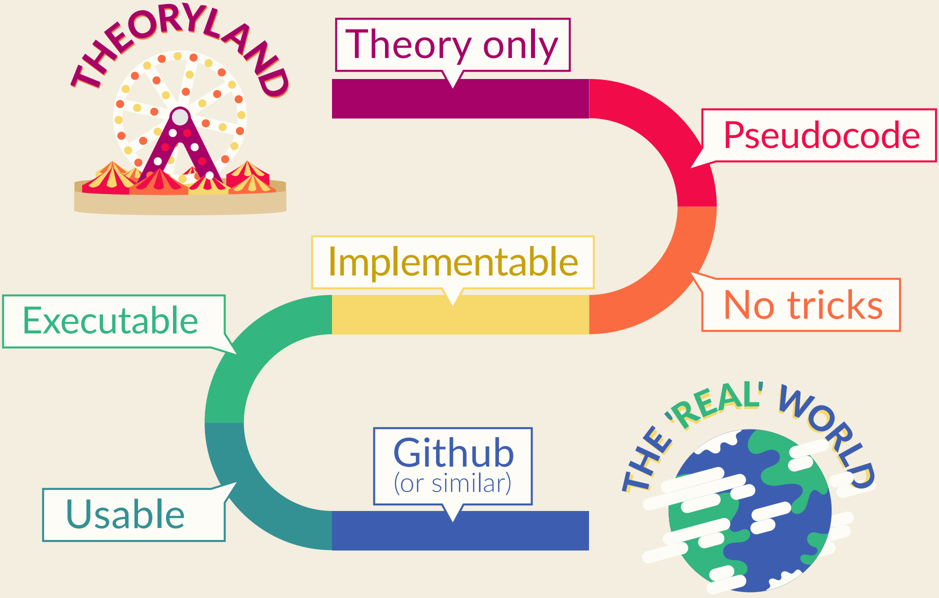
Reduces to
well-known problem

Easy to understand

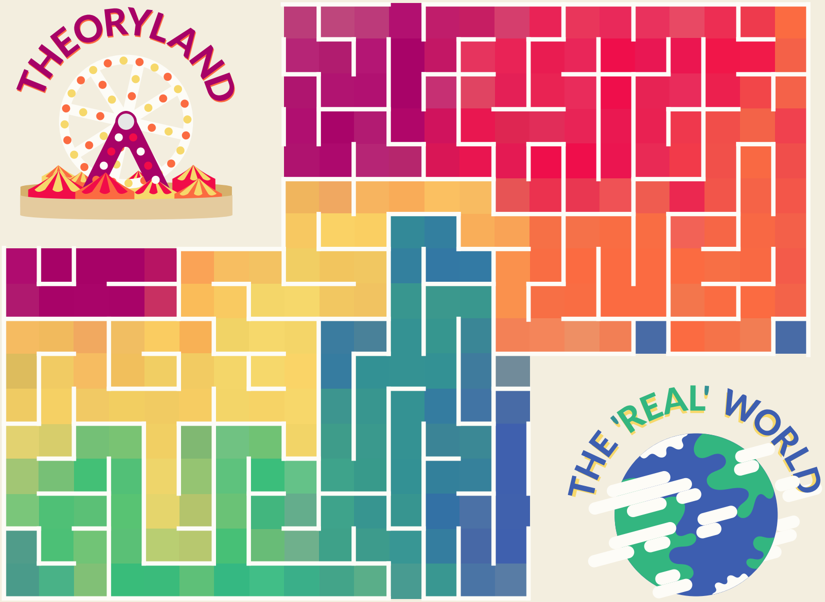
"I just cloned it from github and it ran fine on my data set."



The long and winding road

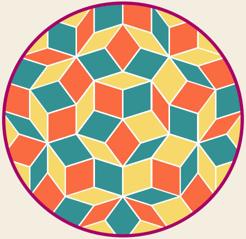
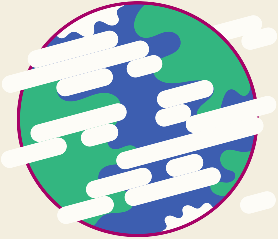


The long and winding ~~road~~



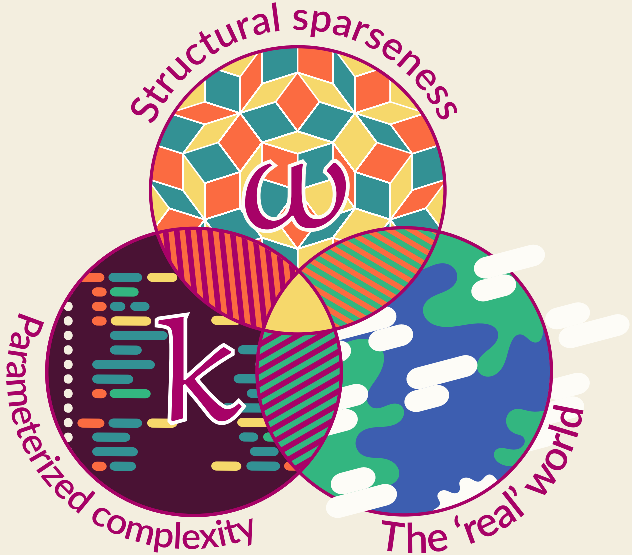
Part II

Background(s)

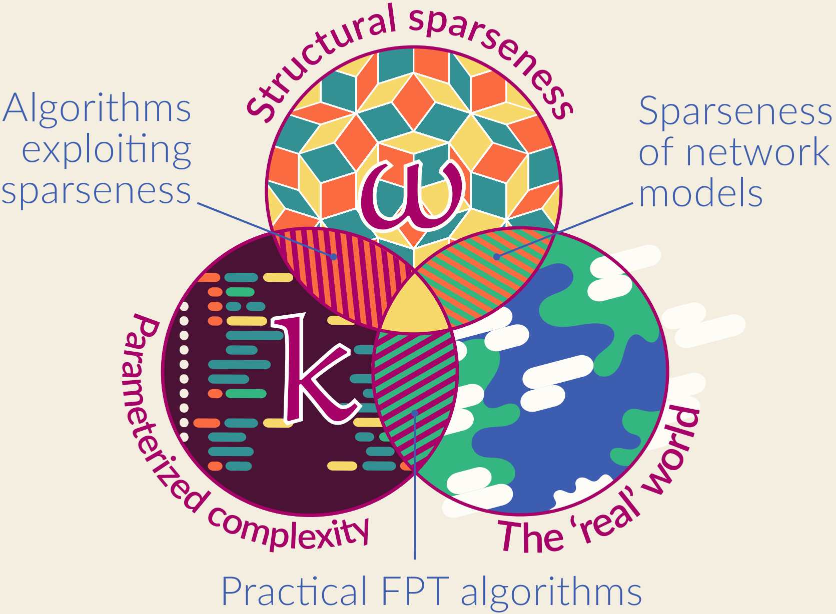
  

Three domains

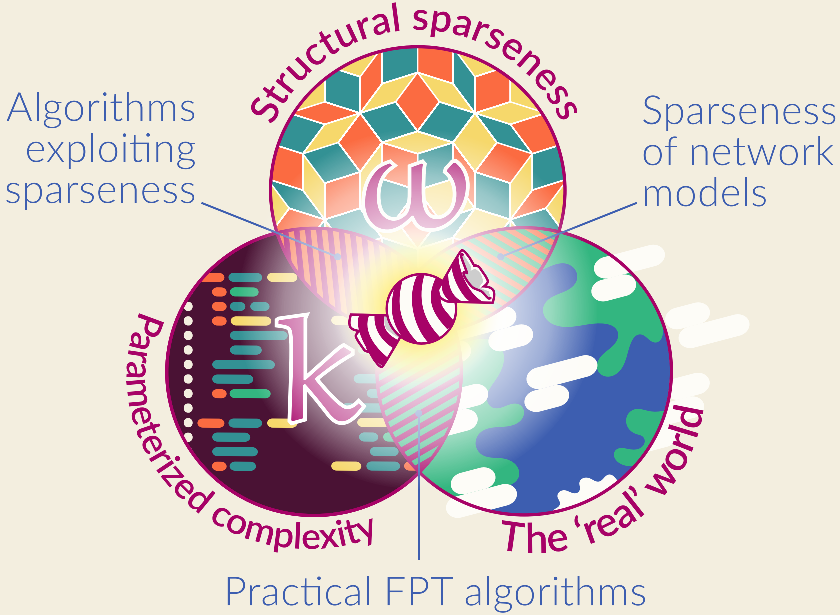
Four intersections



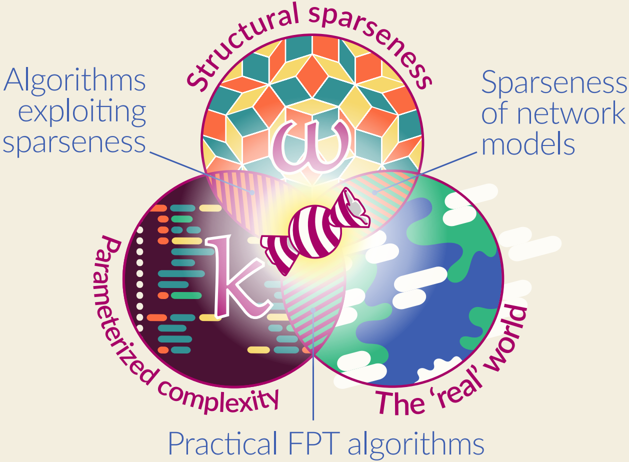
Four intersections

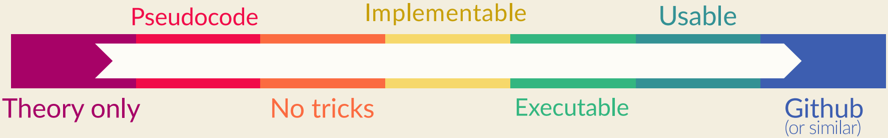


Four intersections

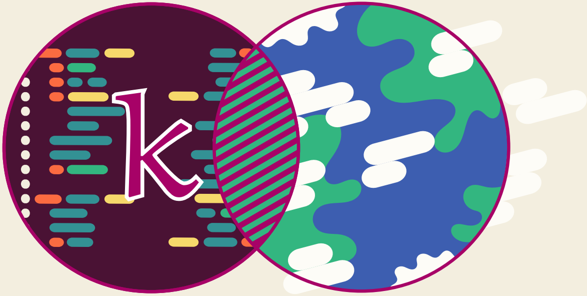


Four intersections





FPT ∩ **real world**



Criticism of trickery in algorithms (ask Mike about the "PTAS industry") was central in proposing parameterized complexity as an alternative[1].

Questions of practicality are encoded in the DNA of the field, and it shows!

[1] Downey R. **Parameterized complexity for the skeptic**. InProc. 18th IEEE Annual Conference on Computational Complexity 2003 Jul 7 (Vol. 132).

FPT ∩ **real world**

"The collection of methods for classifying problems as fixed-parameter tractable, for designing FPT algorithms, for designing better FPT algorithms and **transfering these results to practical implementations**, [...] has developed with surprising vigor."

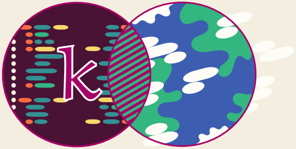
Fellows MR. **Blow-ups, win/win's, and crown rules: Some new directions in FPT**. InInternational Workshop on Graph-Theoretic Concepts in Computer Science 2003 Jun 19 (pp. 1-12). Springer, Berlin, Heidelberg.

Questions of practicality are **raised regularly**, see e.g.

Cheetham J, Dehne F, Rau-Chaplin A, Stege U, Taillon PJ. **Solving large FPT problems on coarse-grained parallel machines**. Journal of Computer and System Sciences. 2003 Dec 31;67(4):691-706.

Langston M. **Practical FPT implementations and applications**. InInternational Workshop on Parameterized and Exact Computation 2004 Sep 14 (pp. 291-291). Springer, Berlin, Heidelberg.

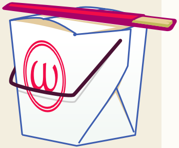
Abu-Khzam FN, Cai S, Egan J, Shaw P, Wang K. **Turbo-Charging Dominating Set with an FPT Subroutine: Further Improvements and Experimental Analysis**. InInternational Conference on Theory and Applications of Models of Computation 2017 Apr 20 (pp. 59-70). Springer, Cham.



FPT $\cap$ **real world**

While many FPT algorithms are impractical (either because of trickery or parameters that are not small in practise), there is a consensus in the field that **practicality matters**.

Parameterized complexity therefore not only gives us a rich theoretical framework, we also find many past attempts of applying said framework to real world problems.



Takeaway:

Parameterized complexity offers the framework, the the institutions, and the culture to investigate the question of practicality.

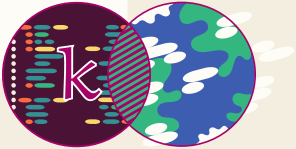
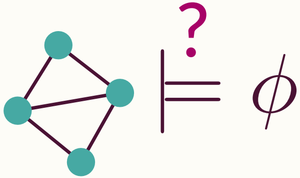


Exhibit A
SEQUOIA

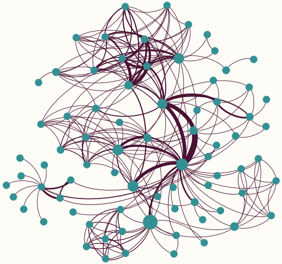
MSO-model checking
against all odds

MSO model checking



Given an MSO_2 formula in the language of (annotated) graphs and a graph, test whether the graph is a model of the formula.

MSO model checking



$$\models^{?} \phi$$

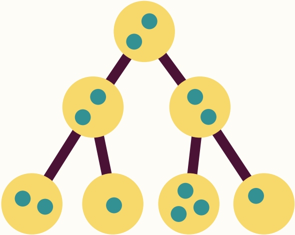
Given an MSO_2 formula in the language of (annotated) graphs and a graph, test whether the graph is a model of the formula.

Not FPT for graphs of moderately unbounded treewidth!

Kreutzer S. **On the Parameterised Intractability of Monadic Second-Order Logic**. InCSL 2009 Sep 7 (Vol. 9, pp. 348-363).

Kreutzer S, Tazari S. **Lower bounds for the complexity of monadic second-order logic**. InLogic in Computer Science (LICS), 2010 25th Annual IEEE Symposium on 2010 Jul 11 (pp. 189-198). IEEE.

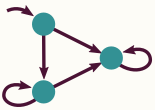
MSO model checking



$$\overset{?}{\models} \phi$$

Given an MSO_2 formula in the language of (annotated) graphs and a graph **of bounded treewidth**, test whether the graph is a model of the formula.

Engineering Courcelle's Theorem



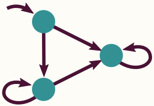
State
explosion

Pseudocode
Implementable
Usable

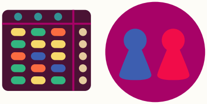
Theory only
No tricks
Executable
Github
(or similar)

Courcelle B. **The monadic second-order logic of graphs. I. Recognizable sets of finite graphs**. Information and computation. 1990 Mar 1;85(1):12-75.

Engineering Courcelle's Theorem



State explosion



DP using Hintikka games

Pseudocode

Implementable

Usable

Theory only

No tricks

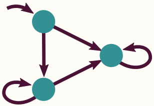
Executable

Github
(or similar)

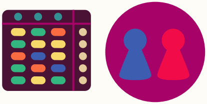
Courcelle B. **The monadic second-order logic of graphs. I. Recognizable sets of finite graphs**. Information and computation. 1990 Mar 1;85(1):12-75.

Kneis J, Langer A, Rossmanith P. **Courcelle's theorem—a game-theoretic approach**. Discrete Optimization. 2011 Nov 30;8(4):568-94.

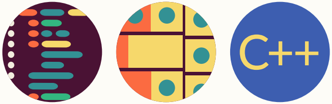
Engineering Courcelle's Theorem



State
explosion



DP using
Hintikka games



Software Engineering
& optimization

Pseudocode — Implementable — Usable

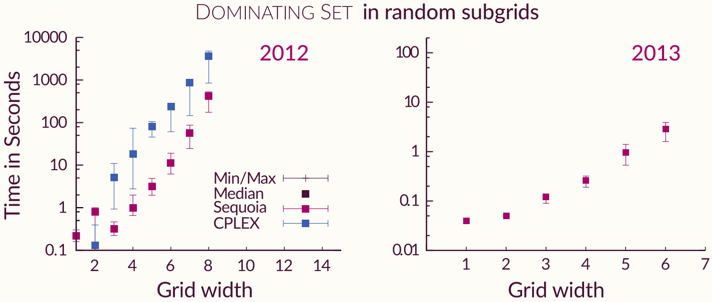
Theory only — No tricks — Executable — Github (or similar)

Courcelle B. **The monadic second-order logic of graphs. I. Recognizable sets of finite graphs.** Information and computation. 1990 Mar 1;85(1):12-75.

Langer A. **Fast algorithms for decomposable graphs.** (Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2013).
https://github.com/sequoia-mso/sequoia-core

Kneis J, Langer A, Rossmanith P. **Courcelle's theorem—a game-theoretic approach.** Discrete Optimization. 2011 Nov 30;8(4):568-94.

Sequoia performance



2013 CONNECTED DOMINATING SET

Instance	n	tw	Sequoia time	Sequoia solution	CPLEX time	CPLEX solution	CPLEX gap
Hannover (small)	673	8	3 min	319	–	327	41%
Hannover (large)	956	9	9 min	376	–	385	42%
Berlin	2599	11	3 h 12 min	1269	–	1342	35%

Langer A. **Fast algorithms for decomposable graphs**.
(Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2013).

Langer A, Reidl F, Rossmanith P, Sikdar S. **Evaluation of an MSO-solver**.
In2012 Proceedings of the Fourteenth Workshop on Algorithm Engineering and Experiments (ALENEX) 2012 Jan 16 (pp. 55-63). Society for Industrial and Applied Mathematics.

Exhibit B

Toboggan

High-throughput FPT

Flow decomposition

AGGACGTAGATAGCTAGCTAATGCTACGATCAGAGGACGTAGATTTATTACCAT

TACCGAATACGAACTAGGATATCGATCGATCAGAGGCCCAATAGGGAATATCCG

TACCGAATACGAACTAGGATATCGATCGATTGATCTATAATAGTAGAATATCCG

Shared segments between DNA/RNA strands create ambiguity in the assembly problem.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**. InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V. **A novel min-cost flow method for estimating transcript expression with RNA-Seq**. BMC bioinformatics. 2013 Apr 10;14(5):S15.

Flow decomposition

AGGACGTAG	ATAGCTAGC	TAATGCTAC	GATCAGAGG	ACGTAGATT	TATTACCAT
TACCGAATA	CGAACTAGG	ATATCGATC	GATCAGAGG	CCCAATAGG	GAATATCCG
TACCGAATA	CGAACTAGG	ATATCGATC	GATTGATCT	ATAATAGTA	GAATATCCG

Shared segments between DNA/RNA strands create ambiguity in the assembly problem.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**. InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V. **A novel min-cost flow method for estimating transcript expression with RNA-Seq**. BMC bioinformatics. 2013 Apr 10;14(5):S15.

Flow decomposition

AGGACGTAG	ATAGCTAGC	TAATGCTAC	GATCAGAGG	ACGTAGATT	TATTACCAT
TACCGAATA	CGAACTAGG	ATATCGATC	GATCAGAGG	CCCAATAGG	GAATATCCG
TACCGAATA	CGAACTAGG	ATATCGATC	GATTGATCT	ATAATAGTA	GAATATCCG

Shared segments between DNA/RNA strands create ambiguity in the assembly problem.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**.
InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V.
A novel min-cost flow method for estimating transcript expression with RNA-Seq.
BMC bioinformatics. 2013 Apr 10;14(5):S15.

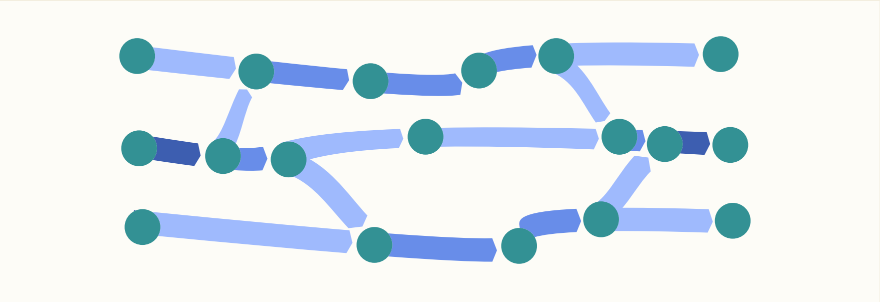
Flow decomposition

Shared segments between DNA/RNA strands create ambiguity in the assembly problem.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**.
InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V.
A novel min-cost flow method for estimating transcript expression with RNA-Seq.
BMC bioinformatics. 2013 Apr 10;14(5):S15.

Flow decomposition

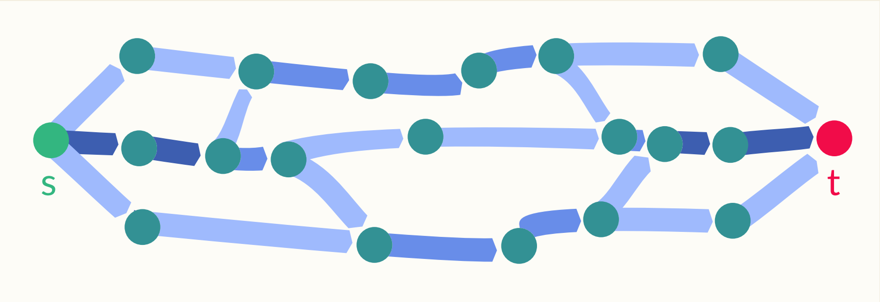


Connecting overlapping segments ('k-mers') and counting their frequencies ('abundance') yields a DAG and a flow.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**.
InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V.
A novel min-cost flow method for estimating transcript expression with RNA-Seq.
BMC bioinformatics. 2013 Apr 10;14(5):S15.

Flow decomposition

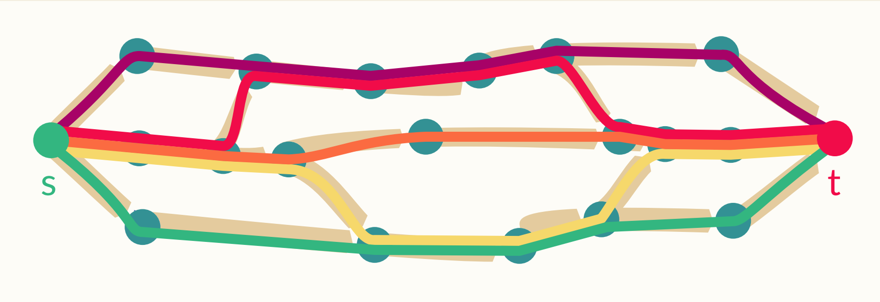


Connecting overlapping segments (‘k-mers’) and counting their frequencies (‘abundance’) yields a DAG and a flow.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**. InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V. **A novel min-cost flow method for estimating transcript expression with RNA-Seq**. BMC bioinformatics. 2013 Apr 10;14(5):S15.

Flow decomposition



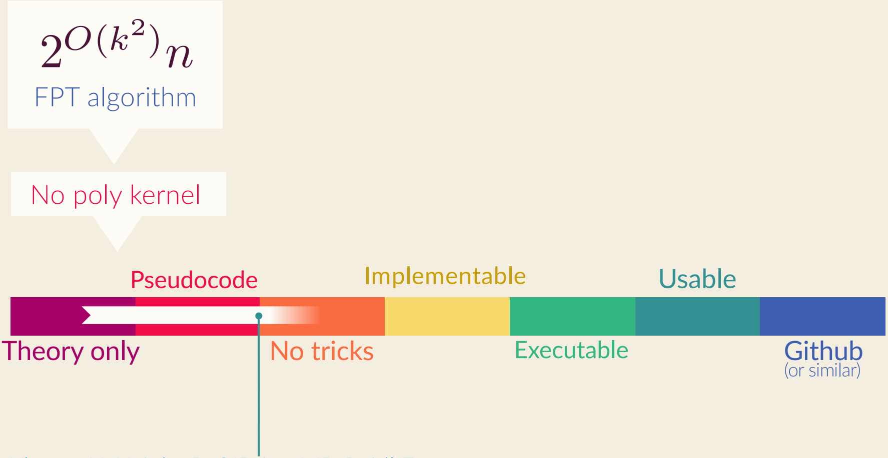
Our task is to split the flow into k **weighted** s-t-paths in order to recover the original DNA/RNA strands.

Hartman T, Hassidim A, Kaplan H, Raz D, Segalov M. **How to split a flow?**.
InINFOCOM, 2012 Proceedings IEEE 2012 Mar 25 (pp. 828-836). IEEE.

Tomescu AI, Kuosmanen A, Rizzi R, Mäkinen V.
A novel min-cost flow method for estimating transcript expression with RNA-Seq.
BMC bioinformatics. 2013 Apr 10;14(5):S15.

f

Engineering flow decomposition

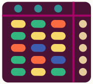


Kloster K, Kuinke P, O'Brien MP, Reidl F,
Sánchez Villaamil F, Sullivan BD, van der Poel A.
A practical fpt algorithm for Flow Decomposition and transcript assembly.
arXiv preprint arXiv:1706.07851. 2017 Jun 23.

Engineering flow decomposition

$$2^{O(k^2)}n$$

FPT algorithm



DP using ILPs*
to encode
constraints

No poly kernel

Pseudocode
Implementable
Usable

Theory only
No tricks
Executable
Github
(or similar)

Kloster K, Kuinke P, O'Brien MP, Reidl F,
Sánchez Villaamil F, Sullivan BD, van der Poel A.
A practical fpt algorithm for Flow Decomposition and transcript assembly.
arXiv preprint arXiv:1706.07851. 2017 Jun 23.

*More precisely: systems of Diophantine equations

Engineering flow decomposition

$2^{O(k^2)}n$

FPT algorithm

No poly kernel

DP using ILPs*
to encode
constraints

Preprocessing, early-out
heuristics and optimization

Pseudocode

Implementable

Usable

Theory only

No tricks

Executable

Github
(or similar)

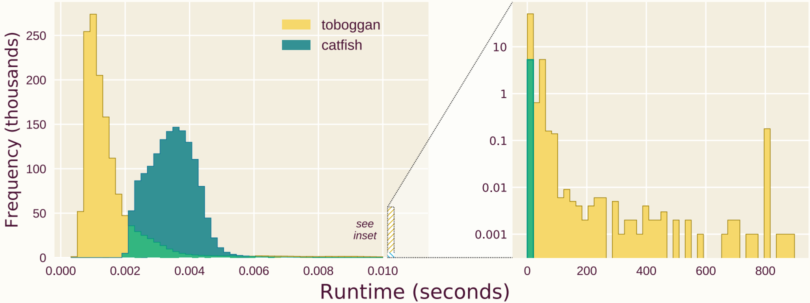
DOI 10.5281/zenodo.821633
https://github.com/TheoryInPractice/toboggan

Kloster K, Kuinke P, O'Brien MP, Reidl F,
Sánchez Villaamil F, Sullivan BD, van der Poel A.
A practical fpt algorithm for Flow Decomposition and transcript assembly.
arXiv preprint arXiv:1706.07851. 2017 Jun 23.

*More precisely: systems of Diophantine equations

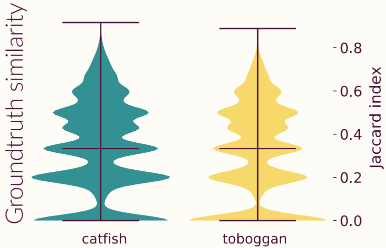
Toboggan vs Catfish

A comparison over ~4 million instances



Groundtruth recovery

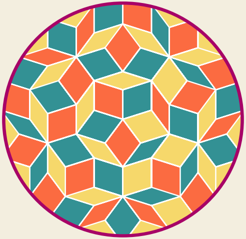
k	instances	Catfish	Toboggan
2	63.2791 %	0.992	0.995
3	22.0775 %	0.967	0.969
4	8.5237 %	0.931	0.930
5	3.4920 %	0.886	0.886
6	1.5375 %	0.830	0.828
7	0.6698 %	0.788	0.780
8	0.2889 %	0.767	0.766
9	0.1241 %	0.740	0.743
10	0.0070 %	0.752	0.802
11	0.0004 %	0.500	0.500
All	100 %	0.973	0.975



Kloster K, Kuinke P, O'Brien MP, Reidl F, Sánchez Villaamil F, Sullivan BD, van der Poel A. **A practical fpt algorithm for Flow Decomposition and transcript assembly.** arXiv preprint arXiv:1706.07851. 2017 Jun 23.

Part III

Once more, with *sparsity*



Elevator pitch: structural sparseness

A **graph measure** is an isomorphism invariant function that maps graphs to $\mathbb{R}^+$.

e.g. density, average degree, clique number, degeneracy treewidth, etc.

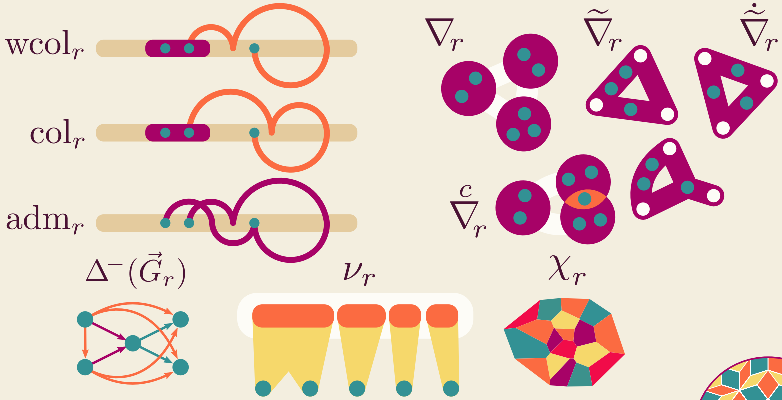
A **parameterised graph measure** is a family of graph measures $(f_r)_{r \in \mathbb{N}_0}$.

A graph class $\mathcal{G}$ is f_r**-bounded** if there exists g s.t.

$$f_r(\mathcal{G}) = \sup_{G \in \mathcal{G}} f_r(G) \leqslant g(r) \quad \text{for all } r.$$

Elevator pitch: structural sparseness

Nešetřil & Ossona de Mendez:
Many notions of f_r-boundedness are equivalent!



We call classes in which these measures are bounded classes of **bounded expansion**

Sparsity ∩ FPT

In “Fixed-Parameter Tractability and Completeness”[1], the first technique Mike and Rod put of forward is the $O(n^3)$ minor-testing routine by Robertson and Seymour.

If practicality in the DNA of the field, the graph minor theorem is its stem cell.

[1] Downey RG, Fellows MR. **Fixed-parameter tractability and completeness**. Cornell University, Mathematical Sciences Institute; 1992.

Sparsity ∩ FPT: **big hammers**

MSO-Model checking in graphs of bounded treewidth

Arnborg S, Lagergren J, Seese D. **Easy problems for tree-decomposable graphs**. Journal of Algorithms. 1991 Jun 1;12(2):308-40.

FO-Model checking in nowhere-dense graphs

Grohe M, Kreutzer S, Siebertz S. **Deciding first-order properties of nowhere dense graphs**. Journal of the ACM (JACM). 2017 Jun 16;64(3):17.

Meta-Kernelization in sparse classes

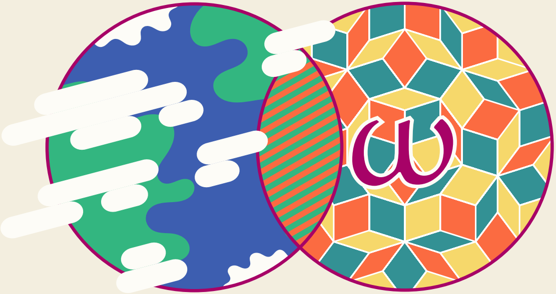
Bodlaender HL, Fomin FV, Lokshtanov D, Penninkx E, Saurabh S, Thilikos DM. **(Meta) kernelization**. InFoundations of Computer Science, 2009. FOCS'09. 50th Annual IEEE Symposium on 2009 Oct 25 (pp. 629-638). IEEE.

Fomin FV, Lokshtanov D, Saurabh S, Thilikos DM. **Bidimensionality and kernels**. InProceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms 2010 Jan 17 (pp. 503-510). Society for Industrial and Applied Mathematics.

Gajarský J, Hliněný P, Obdržálek J, Ordyniak S, Reidl F, Rossmanith P, Sánchez Villaamil F, Sikdar S. **Kernelization using structural parameters on sparse graph classes**. Journal of Computer and System Sciences. 2017 Mar 31;84:219-42.

Kim EJ, Langer A, Paul C, Reidl F, Rossmanith P, Sau I, Sikdar S. **Linear kernels and single-exponential algorithms via protrusion decompositions**. ACM Transactions on Algorithms (TALG). 2016 Feb 12;12(2):21.

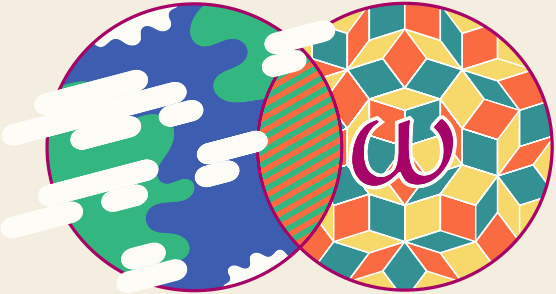
Real world ∩ sparsity



The big question:

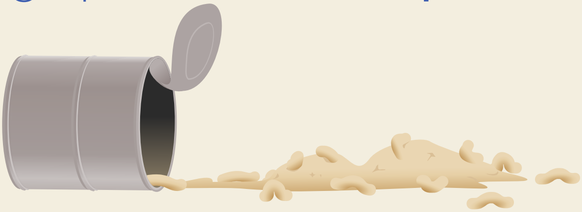
Are real-world graphs (*networks*) **sparse**?

Real world ∩ sparsity



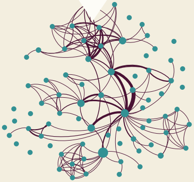
The big question:

Are real-world graphs (*networks*) **sparse**?



Graph classes vs. networks

Network instances



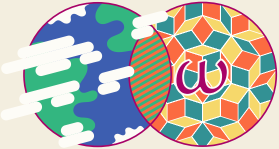
Mathematical Theory

Theorem.

$$Pr[\|G\| \geq \xi k] \leq \left(\frac{e\beta D^2}{2n\xi k e^{D^2/2n}}\right)^{\xi k}$$

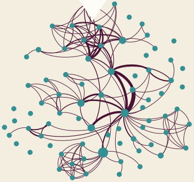
???

...



Graph classes vs. networks

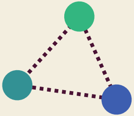
Network instances



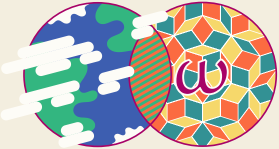
Mathematical Theory

Theorem.

$$Pr[\|G\| \geq \xi k] \leq \left(\frac{e\beta D^2}{2n\xi k e^{D^2/2n}}\right)^{\xi k}$$

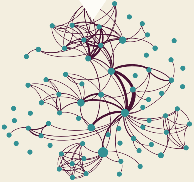


Network model



Graph classes vs. networks

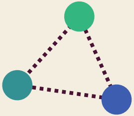
Network instances



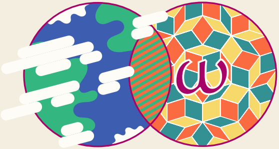
Mathematical Theory

Theorem.

$$Pr[\|G\| \geq \xi k] \leq \left(\frac{e\beta D^2}{2n\xi k e^{D^2/2n}}\right)^{\xi k}$$



Network model



Structurally sparse random graphs

Sparse Erdős–Rényi graphs are structurally sparse

Nešetřil J, de Mendez PO, Wood DR.
Characterisations and examples of graph classes with bounded expansion.
European Journal of Combinatorics. 2012 Apr 30;33(3):350-73.

Random graphs with fixed degree distribution that have quickly vanishing tails are structurally sparse

Demaine ED, Reidl F, Rossmanith P, Sánchez Villaamil F, Sikdar S, Sullivan BD.
Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs.
arXiv preprint arXiv:1406.2587. 2014 Jun 10.

Random intersection graphs (in certain regimes) are structurally sparse.

Farrell M, Goodrich TD, Lemons N, Reidl F, Sánchez Villaamil F, Sullivan BD.
Hyperbolicity, degeneracy, and expansion of random intersection graphs.
InInternational Workshop on Algorithms and Models for the Web-Graph 2015 Dec 10 (pp. 29-41). Springer, Cham.

Structurally sparse random graphs

Sparse Erdős–Rényi graphs are structurally sparse

Nešetřil J, de Mendez PO, Wood DR.
Characterisations and examples of graph classes with bounded expansion.
European Journal of Combinatorics. 2012 Apr 30;33(3):350-73.

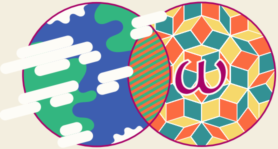
Random graphs with fixed degree distribution that have quickly vanishing tails are structurally sparse

Demaine ED, Reidl F, Rossmanith P, Sánchez Villaamil F, Sikdar S, Sullivan BD.
Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs.
arXiv preprint arXiv:1406.2587. 2014 Jun 10.

Random intersection graphs (in certain regimes) are structurally sparse.

Farrell M, Goodrich TD, Lemons N, Reidl F, Sánchez Villaamil F, Sullivan BD.
Hyperbolicity, degeneracy, and expansion of random intersection graphs.
InInternational Workshop on Algorithms and Models for the Web-Graph 2015 Dec 10 (pp. 29-41). Springer, Cham.

Are these models in these regimes representative of real world graphs?



Structurally sparse random graphs

Sparse Erdős–Rényi graphs are structurally sparse

Nešetřil J, de Mendez PO, Wood DR.
Characterisations and examples of graph classes with bounded expansion.
European Journal of Combinatorics. 2012 Apr 30;33(3):350-73.

Random graphs with fixed degree distribution that have quickly vanishing tails are structurally sparse

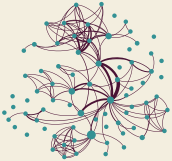
Demaine ED, Reidl F, Rossmanith P, Sánchez Villaamil F, Sikdar S, Sullivan BD.
Structural sparsity of complex networks: Bounded expansion in random models and real-world graphs.
arXiv preprint arXiv:1406.2587. 2014 Jun 10.

Random intersection graphs (in certain regimes) are structurally sparse.

Farrell M, Goodrich TD, Lemons N, Reidl F, Sánchez Villaamil F, Sullivan BD.
Hyperbolicity, degeneracy, and expansion of random intersection graphs.
InInternational Workshop on Algorithms and Models for the Web-Graph 2015 Dec 10 (pp. 29-41). Springer, Cham.

Are these models in these regimes representative of real world graphs?

Real structural sparseness



28

Statistical test of degree distribution

Tail looks subcubic

6

Cannot decide

Tail looks supercubic

22

Compare density to Chung-Lu w/ same degree distribution

Denser than Chung-Lu

8

Cannot decide

7 of those are probably sparse

Sparser than Chung-Lu

14

Structurally sparse!

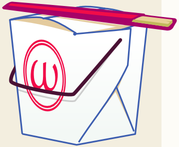
Reidl F. **Structural sparseness and complex networks.**
(Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

Real world ∩ sparsity

All experiments so far as well as properties of several important network models point towards complex networks being structurally sparse.

Some of these results directly contradict some widely held assumptions about networks!

There are good reasons to dismiss these assumptions.
Don't ask me about that unless you want a half-hour rant.



Takeaway:
Many real-world networks are structurally sparse.

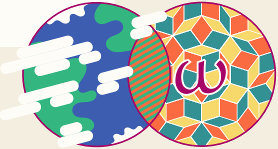


Exhibit C

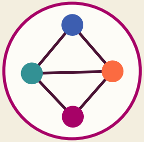
CONCUSS

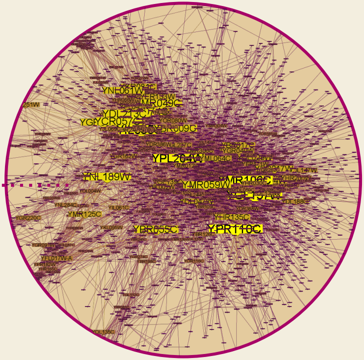
Combatting Network Complexity
Using Structural Sparsity

Motif counting using low-treedepth colourings

Motif-counting

We want to count the number of times a given **motif graph**





appears in a larger host graph.

Motifs that appear more often than expected probably play an important role in the network

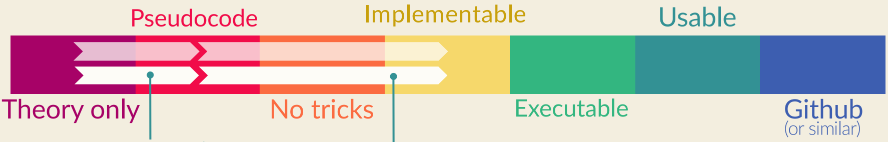
Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. **Network motifs: simple building blocks of complex networks**. Science. 2002 Oct 25;298(5594):824-7.

Ribeiro P, Silva F, Kaiser M. **Strategies for network motifs discovery**. InE-Science, 2009. e-Science'09. Fifth IEEE International Conference on 2009 Dec 9 (pp. 80-87). IEEE.

Engineering motif-counting

$$f(h) \cdot 2^{O(h^2)} n$$

$$f(h) \cdot h^{O(h)} n$$

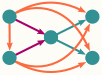


Nešetřil J, De Mendez PO.
Sparsity: graphs, structures, and algorithms.
Springer Science & Business Media; 2012 Apr 24.

Demaine ED, Reidl F, Rossmanith P,
Sánchez Villaamil F, Sikdar S, Sullivan BD.
Structural sparsity of complex networks:
Bounded expansion in random models and real-world graphs.
arXiv preprint arXiv:1406.2587. 2014 Jun 10.

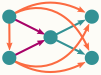
Engineering motif-counting

tf-augmentations



Absolutely impractical

dtf-augmentations



Test colouring after each step

Pseudocode Implementable Usable

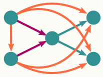
Theory only No tricks Executable Github
(or similar)

Nešetřil J, De Mendez PO.
Sparsity: graphs, structures, and algorithms.
Springer Science & Business Media; 2012 Apr 24.

Reidl F. **Structural sparseness and complex networks.**
(Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

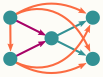
Engineering motif-counting

tf-augmentations



Absolutely impractical

dtf-augmentations



Test colouring after each step

Good engineering & heuristical improvements

Pseudocode

Implementable

Usable

Theory only

No tricks

Executable

Github (or similar)

Nešetřil J, De Mendez PO.
Sparsity: graphs, structures, and algorithms.
Springer Science & Business Media; 2012 Apr 24.

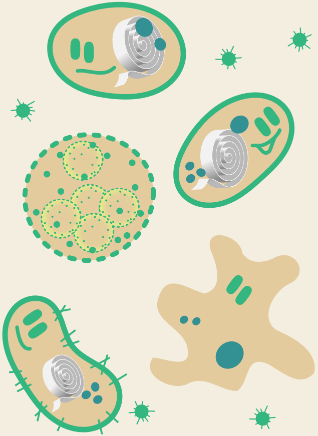
Reidl F. **Structural sparseness and complex networks.**
(Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

Theory in Practice Group (NCSU)
with great help from students
Clayton G. Hobbs & Brandon Mork
https://github.com/TheoryInPractice/CONCUSS

Exhibit D
CATLAS

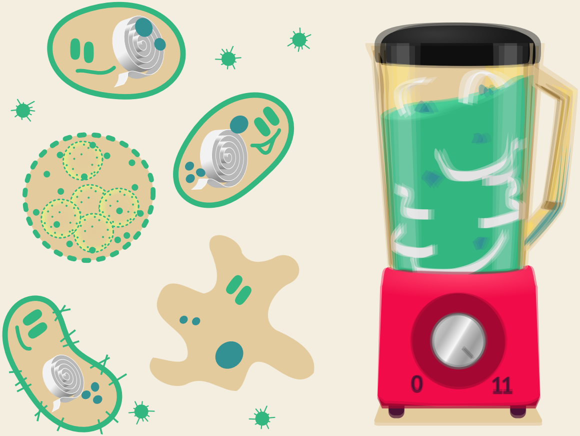
Metagenome exploration
using hierarchical domination
of de-Bruijn graphs

Metagenomics



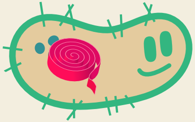
Metagenomics

Metagenomics

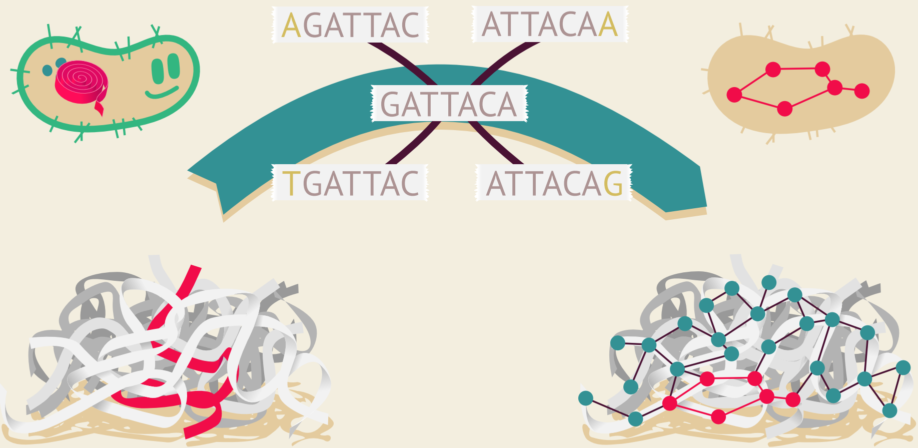


Metagenomics

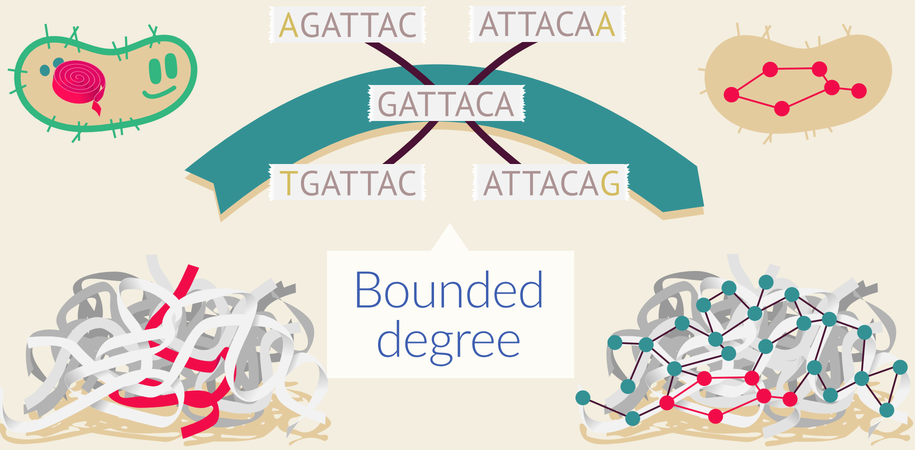
De-Bruijn graphs



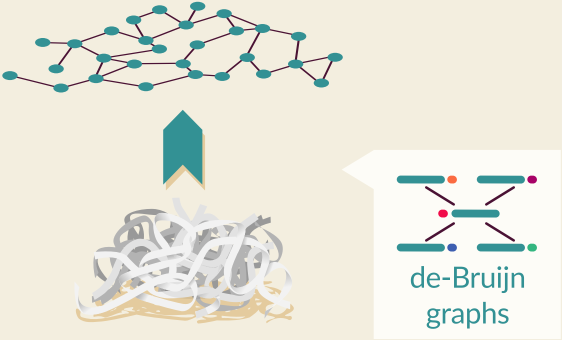
De-Bruijn graphs



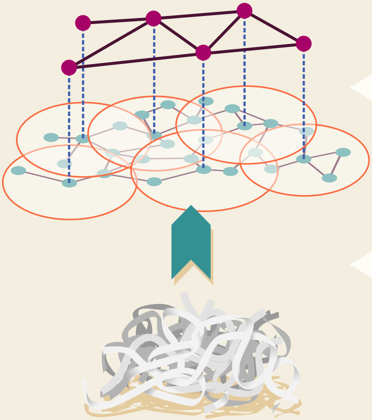
De-Bruijn graphs

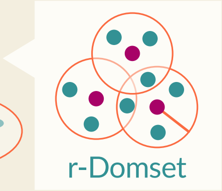


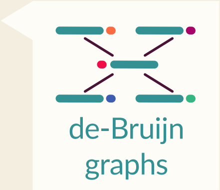
CATLAS Overview



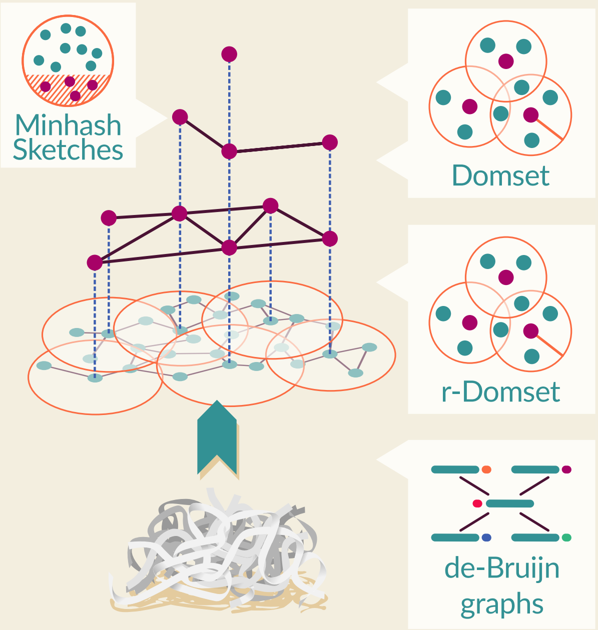
CATLAS Overview



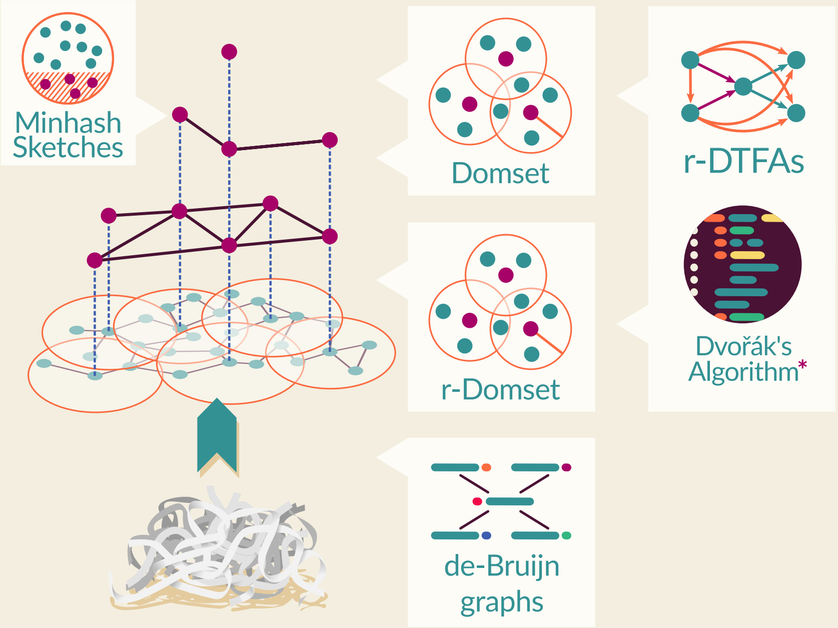




CATLAS Overview



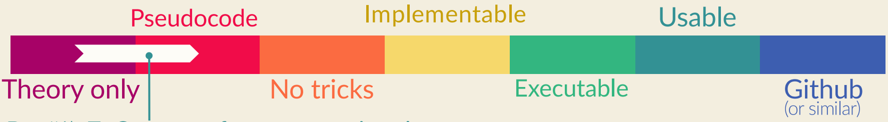
CATLAS Overview



Engineering Dvořák's algorithm

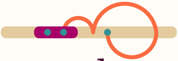
wcol_{2r}

Approximation is terrible in practise



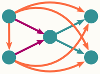
Dvořák Z. **Constant-factor approximation of the domination number in sparse graphs**. European Journal of Combinatorics. 2013 Jul 31;34(5):833-40.

Engineering Dvořák's algorithm



wcol_{2r}

Approximation is terrible in practise



$\Delta^{-}(\vec{G}_{2r})$

Approximation is terrible in practise

Pseudocode

Implementable

Usable

Theory only

No tricks

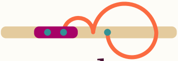
Executable

Github
(or similar)

Dvořák Z. **Constant-factor approximation of the domination number in sparse graphs.** European Journal of Combinatorics. 2013 Jul 31;34(5):833-40.

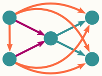
Reidl F. **Structural sparseness and complex networks.** (Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

Engineering Dvořák's algorithm



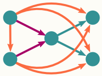
wcol_{2r}

Approximation is terrible in practise



$\Delta^{-}(\vec{G}_{2r})$

Approximation is terrible in practise



$\Delta^{-}(\vec{G}_{r})$

Approximation is tunable (heuristic)

Pseudocode | Implementable | Usable

Theory only | No tricks | Executable | Github (or similar)

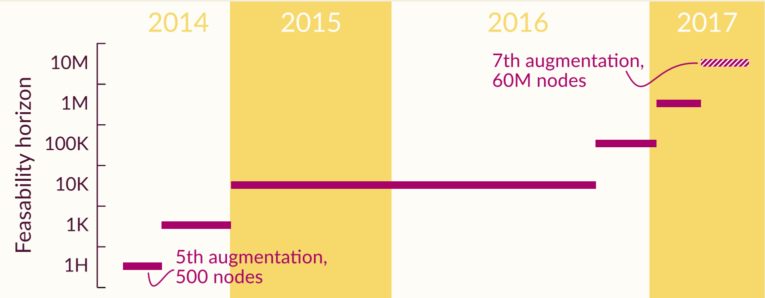
Dvořák Z. **Constant-factor approximation of the domination number in sparse graphs.** European Journal of Combinatorics. 2013 Jul 31;34(5):833-40.

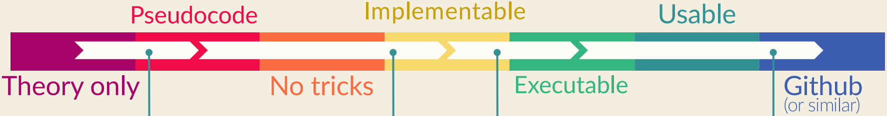
Reidl F. **Structural sparseness and complex networks.** (Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

Barnraising for Data-Intensive Discovery at MDI Biological Laboratory in Salisbury Cove, Maine.

Ongoing collaboration with **Theory in Practice Group** (NCSU) and **Lab for Data Intensive Biology** (UC Davis)
https://github.com/spacegraphcats/spacegraphcats

Orders of magnitude





Dvořák Z. **Constant-factor approximation of the domination number in sparse graphs.** European Journal of Combinatorics. 2013 Jul 31;34(5):833-40.

Reidl F. **Structural sparseness and complex networks.** (Doctoral dissertation, Dissertation, Aachen, Techn. Hochsch., 2015).

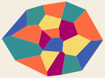
Barnraising for Data-Intensive Discovery at MDI Biological Laboratory in Salisbury Cove, Maine.

Ongoing collaboration with **Theory in Practice Group** (NCSU) and **Lab for Data Intensive Biology** (UC Davis)
https://github.com/spacegraphcats/spacegraphcats

Part IV

Future research

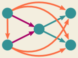
Graph measures



Measures often NP-hard even for very low r

Muzi I, O'Brien MP, Reidl F, Sullivan BD. **Being even slightly shallow makes life hard**. Accepted at MFCS'17

Z. Dvořák. **Asymptotical Structure of Combinatorial Objects**. (PhD thesis, Charles University, Faculty of Mathematics and Physics, 2007).



- Polynomial-time computable measure (for some r > 0)?
- Find fast exact algorithms or (good) approximations
- Find new bounded expansion characterisations
- Further map out relationship between measures

- Find ‘typical’ values of measures in networks
- Design heuristic taylored to real-world instances

Template: Bodlaender's Algorithm, treewidth heuristics

Meta-theorems & algorithms

Approximation in bounded expansion classes

Polynomial-time algorithms in sparse classes

Hand-design algorithms for selected problems

Template: FPT running time/kernel size races

FO model-checking ‘without tricks’

Identify suitable FO-fragment

Make logic approachable for normal folks

Implement more algorithms (student projects!)

Template: Courcelle's Theorem vs. hand-crafted DP

A sparse theory of density

A theory of dense graphs with underlying sparse structure

A matching algorithmic meta-theory

Template: Treewidth/Rankwidth, Sparsity programme, FO-model checking in nowhere dense graphs

Hand-crafted FPT algorithms in such graphs

Apply to biocomp problems in e.g protein-protein interaction networks

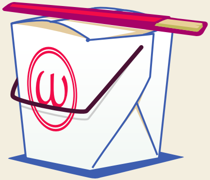
Template: None yet?

Takeaway

The combination of structural sparseness and parameterized algorithms has **the potential to deliver practical algorithms**.

The challenge of making theoretical results applicable in practise usually generates **theoretical follow-up questions**.

Aiming for practicality imposes restraints (no tricks!), but that is **not necessarily a bad thing**.



THANKS!

Questions?

A big thanks to my co-authors:

Fernando Sánchez Villaamil
Somnath Sikdar
Peter Rossmanith
Blair D. Sullivan
Alexander Langer
Michael P. O'Brien
Jakub Gajarský
Pål Grønås Drange
Jan Obdrzálek
Sebastian Ordyniak
Timothy Goodrich
Petr Hlinený
Andrew van der Poel
Erik D. Demaine

Eun Jung Kim
Magnus Wahlström
Christophe Paul
Gregory Gutin
Ignasi Sau
Michal Pilipczuk
Nathan Lemons
Philipp Kuinke
Stephan Kreutzer
Saket Saurabh
Alex J. Chin
Fedor V. Fomin
Marcin Pilipczuk

Markus S. Dregi
Martin L. Demaine
Matthew Farrell
Aaron B. Adcock
Daniel Lokshtanov
Irene Muzi
Jan Dreier
Konstantinos Stavropoulos
Kyle Kloster
Sebastian Siebertz
Li-Hsuan Chen
M. S. Ramanujan
Meirav Zehavi