On the Proof Complexity of Combinatorial Principles

(the role of kernelization)

Gabriel Istrate, University of Bucharest

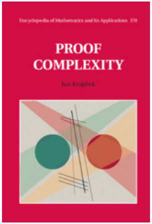
gabriel.istrate@unibuc.ro

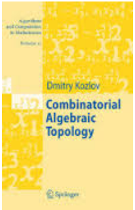


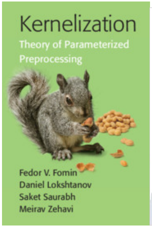


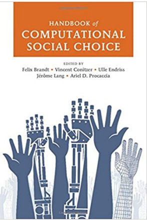
With Adrian Crãciun (Timişoara), James Aisenberg, Sam Buss (both San Diego), Maria-Luisa Bonet (Barcelona), Cosmin Bonchiş(Timişoara).

Take-home message





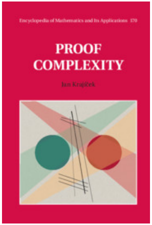




- Motivation: separating complexities of Frege and extended Frege proofs.
- Answer: notions from parameterized complexity can help us obtain efficient Frege proofs.
- Applications: various statements, including ones from combinatorial topology and computational social choice.

Caution: Emphasis on "the story", rather than technical details.

Reminder: Propositional proof complexity



- Proof systems for propositional unsatisfiability, e.g. resolution
- C or x, D or $\overline{x} \rightarrow (C$ or $D)$; $x, \overline{x} \rightarrow \square$.
- Complexity $=$ minimum length of a proof.

EXAMPLE: Pigeonhole principle tautologies PHP^n_{n-1} have $2^{\Omega(n)}$ resolution complexity.

$\boxed{n \text{ pigeons in } n-1 \text{ holes} \Rightarrow \text{at least two pigeons in same hole !}}$

Proof complexity of the pigeonhole principle

- Pigeonhole formula(s): PHP_n^{n-1}
- $X_{i,j} = 1$ "pigeon i goes to hole j".
- $X_{i,1}$ or $X_{i,2}$ or $\ldots$ or $X_{i,n-1}$, $1 \leq i \leq n$ (each pigeon goes to (at least) one hole)
- $\overline{X_{k,j}}$ or $\overline{X_{l,j}}$ (pigeons k and l do not go together to hole j).
- Resolution complexity: $2^{\Omega(n)}$! (Haken)

Buss (J. Symb. Logic): PHP_n^{n-1} has poly-size Frege proofs.

Frege proofs?

- @ boundaries of proof complexity: Frege proofs. For concreteness [Hilbert Ackermann]
  - propositional variables $p_1, p_2, \ldots.$, connectives $\neg$, or .
  - Axiom schematas:
    1. $\neg(A \text{ or } A) \text{ or } A$
    2. $\neg A \text{ or } (A \text{ or } B)$
    3. $\neg(A \text{ or } B) \text{ or } (B \text{ or } A)$
    4. $\neg(\neg A \text{ or } B) \text{ or } (\neg(C \text{ or } A) \text{ or } (C \text{ or } B))$
  - Rule: From A and $\neg A$ or B derive B.
- Other systems, sequent calculus (LK), etc.

All Frege proof systems equivalent (polynomially simulate eachother) S.A. Cook,R. Reckhow. "The relative efficiency of propositional proof systems." J. Symb. Logic 44.1(1979):36-50.

Frege versus extended Frege

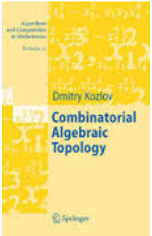
- extended Frege: Frege + variable substitutions $X \leftrightarrow \Phi(\overline{Y})$. Proves same formulas, perhaps more efficiently.

Open: Is extended Frege more powerful than Frege ?

Bonet, M.L., S. Buss, T. Pitassi. "Are there hard examples for Frege systems?." Feasible Mathematics II. Birkhauser, 1995. 30-56.

- Most natural formulas: (quasi)polynomial $(2^{log(n)^{O(1)}})$ Frege proofs.
- Some examples: "$(AB = I) \Rightarrow (BA = I)$" tautologies [Hrubeš, Tzameret CCC'2009], Paris-Harrington tautologies [Carlucci, Galesi, Lauria. CCC, 2011], Frankl Theorem [Buss et al. 2014].

Wishful thinking (around 2014)



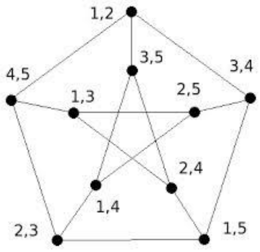
Perhaps translating into SAT a statement that is (mathematically) hard to prove yields a natural candidate for the separation.

- (Martin Kneser, Jaresbericht DMV 1955): Let $n \geq 2k-1 \geq 1$. Let $c : \binom{n}{k} \to [n-2k+1]$. Then there exist two disjoint sets A and B with $c(A) = c(B)$.
- $k = 1$: Pigeonhole principle !
- $k = 2, 3$: combinatorial proofs (Stahl, Garey & Johnson)
- $k \geq 4$: proved in 1977 (Lovász) using Algebraic Topology.

Combinatorial proofs (Matousek, Ziegler). "hide" Alg. Topology No "purely combinatorial" proof (was) known.

Kneser's Conjecture (II)

- the chromatic number of a certain graph $Kn_{n,k}$ (at least) $n - 2k + 2$. (exact value)
- Vertices: $\binom{n}{k}$. Edges: disjoint sets.
- E.g. $k = 2$, $n = 5$: Petersen's graph has chromatic number (at least) three.



First results (paper @ SAT conference)

- naïve encoding $X_{A,k} = TRUE$ iff A colored with color k. Extends encoding of PHP
- $X_{A,1}$ or $X_{A,2}$ or $\ldots$ or $X_{A,n-2k+1}$ "every set is colored with (at least) one color"
- $\overline{X_{A,j}}$ or $\overline{X_{B,j}}$ $(A \cap B = \emptyset)$ "no two disjoint sets are colored with the same color"
- $k = 1$: PHP (studied by Buss).

- $Kneser_n^k$ reduces to $Kneser_{n+2}^{k+1}$.
- $k = 2$: poly-size Frege proofs.
- $k = 3$: poly-size extended Frege proofs.

(paper @ ICALP ⇒ Information and Computation)

- For every $k \geq 3$ $Kneser_k$ poly-size extended Frege proofs, quasi-poly-size Frege proofs.
- For every $k \geq 1$ can reduce verification of $Kneser_k$ to that of a finite number of examples.

For every fixed k, $Kneser_k$ has combinatorial proofs.

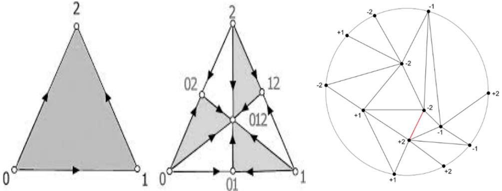
- Mathematically: Kneser follows from octahedral Tucker lemma (algebraic topology, exponential-size objects).
  - "Miniaturization" of this principle: truncated octahedral Tucker lemma.
  - class of propositional formulae, implies Kneser; candidates for separation.

Discrete version of Borsuk-Ulam: Octahedral Tucker's lemma

- Antipodally Symmetric Triangulation T of the n-ball. Barycentric subdivision, one vertex for each face

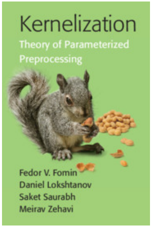
> - For any labeling of T with vertices from $\{\pm 1, \ldots, \pm(n-1)\}$ antipodal on the boundary there exist two adjacent vertices $v \sim w$ with $c(v) = -c(w)$.

- Intuition: no continuous (a.k.a simplicial) antipodal map from the n-ball to the n-sphere.



Second surprise: reverse-engineering proof of Kneser

(second paper @ICALP)



- For every $k \geq 1$ can "reduce" verifying an infinite number of examples to a finite number.
- Behind this type of reduction: kernelization.

Algorithmics, technique for preprocessing individual instances of a combinatorial problem.

Two-minute parameterized complexity

- Many problems in NP parameterized: instance size n, parameter k.
- Can get: complexity $O(n^k)$.
- Parameterized complexity: want complexity $O(f(k) \cdot poly(n))$.
- Kernelization: reduce instance (x, k) to "kernel instance" (x', k'), s.t. $(x, k) \in L$ iff $(x', k') \in L$ and

  $$\boxed{|x'|, k' \leq g(k) \text{ for some computable } g.}$$

- data reduction: algorithm A, maps (x, k) to (x', k') s.t. $(x, k) \in L$ iff $(x', k') \in L$ and $|x'| \leq |x|$, $k' \leq k$. Only for $|x| > g(k)$.
- algorithm: data reduction + bruteforce kernel instances.

- E.g. Vertex Cover: Given graph G and integer k, decide whether G has VC of size at most k. set of vertices that covers all edges.

Rule 1: v isolated vertex in G. G has VC of size k iff $G \setminus \{v\}$ has VC of size k.

Rule 2: v vertex in G, $deg(v) > k$. G has VC of size k iff $G \setminus (\{v\} \cup N(v))$ has VC of size $k - 1$.

THEOREM (parameterized complexity, informal): If G is a graph with more than k^2 vertices then one of Rules 1 and 2 can be applied.

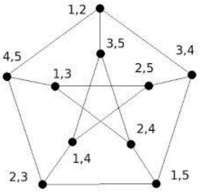
Main idea

- "Negative" instance (x, k) of parameterized problem in NP maps "canonically" to formula $\Phi(x, k) \in \overline{SAT}$.

- If Π_i proof for soundness of the i'th reduction step $(x_i, k_i) = A(x_{i-1}, k_{i-1})$ and Π_{m+1} is a "brute force proof of unsatisfiability" for the kernel instance then one can prove $\Phi(x, k) \in \overline{SAT}$ by "concatenating" $\Pi_1, \ldots, \Pi_m$ and Π_{m+1}.
- This (usually) yields extended Frege proofs.
- For Frege proofs need $m = O(\log n)$ $(m = O(1))$.

Main (meta)Theorem

- Somewhat too complicated to state precisely.
- If soundness of reduction rules can be witnessed efficiently in Frege, the length of reduction chains is $O(1)$ then unsatisfiable formulas $\Phi(x, k)$ have polynomial size Frege proofs.
- If soundness of reduction rules can be witnessed efficiently in Frege, the length of reduction chains is $O(log(|\Phi(x, k)|))$ then unsatisfiable formulas $\Phi(x, k)$ have quasipolynomial Frege proofs.
- otherwise we normally get polynomial size extended Frege proofs.



- Note: inner cycle already chromatic # 3.
- $A \in \binom{n}{k}$ stable if it doesn't contain consecutive elements i, $i+1$ (including $n, 1$).
- Schrijver's Thm.: Chromatic number of stable Kneser graph is $n - 2k + 2$. A. Schrijver. Vertex-critical Subgraphs of Kneser-graphs. N. Arch. Wiskunde XXVI (1978).

THEOREM: For every $k \geq 1$ Schrijver's theorem has quasi-poly size Frege proofs (poly-size Frege)

- Proof idea: data reduction of length $O(\log n)$.

Critical ingredient

We show that $\Theta(n)$ color classes c are star-shaped, i.e. sets colored with color c have an element in common. Need version of Talbot (Intersecting families of separated sets. Journal of the London Mathematical Society, 68(1):37-51, 2003) that can be simulated propositionally:

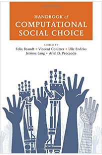
Theorem
If $\mathcal{C}$ is a color class that is not star-shaped then $|\mathcal{C}| \leq k^2 \cdot \binom{n+k-1}{k-2}$.

Thus if there were a $n-2k+1$ coloring c of $SKn_{n,k}$ then we could drop $r = \Theta(n)$ elements of $\{1, 2, \ldots, n\}$ and equally many colors, and reduce the problem to showing that $\chi(SKn_{n-r,k}) > n - r - 2k + 1$.

A Couple of Applications to Proof Complexity

- classical (ad-hoc) kernelization for VertexCover $\Rightarrow$ for every fixed k, negative instances of VC with parameter k have poly-size Frege proofs.
- crown decomposition for DualColoring $\Rightarrow$ negative instances of DualColoring with parameter k poly-size Frege proofs.
- improved (ad-hoc) kernelization for Edge Clique Cover $\Rightarrow$ negative instances (G,k) of Edge Clique Cover have extended Frege proofs of poly size and Frege proofs of quasipoly size.
- sunflower lemma-based kernelization of d-HittingSet $\Rightarrow$ negative instances of d-HittingSet with parameter k extended Frege proofs of poly size.

Applications: Computational Social Choice



- Arrow, Gibbard-Satterthwaite: Fundamental impossibility results on ranking m objects by n agents.
- Tang & Lin (Artificial Intelligence, 2009): Arrow's Theorem has computer-assisted propositional proofs by reducing the general case to the case $n = 2, m = 3$. Similar results (2008) for the Gibbard-Satterthwaite theorem.
- Their proofs: data reductions of length $\Theta(n + m)$.

We give: data reductions of length $O(n)$. Consequently, formulas $Arrow_{m,n}$, $GS_{m,n}$ have (i). quasipoly size Frege proofs (ii). poly size Frege proofs for fixed n.

Conclusions

- Theoretically interesting connections between different areas.
- Work in progress:
  - Adapt this program to other techniques from parameterized complexity, e.g. iterative compression.
  - Adapt this program to other proof systems, e.g. SPR^{-} (Heule,Kiesl & Biere HVC'17, J. Autom. Reasoning '19, Buss & Thapen SAT'19).
  - Proof system that only preserves equisatisfiability, not equivalence
  - Proof complexity for statements in judgment aggregation.
- Proof complexity lower bounds for hard problems in parameterized complexity ?

Where I Would Like to Go

- (Combinatorial) Algebraic Topology: works with exponential size objects.
- Proof Complexity: Cook-Reckow. A proof should be:
  - verifiable in polynomial time.
  - complete.

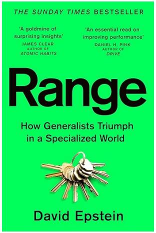
What about non-complete/non-constructive proof systems?

- implicit proofs (Krajicek)
- oracle proof systems (Cook)

On the Combinatorial Algebraic Topology side: e.g. R. Živaljevič. User's guide to equivariant methods in Combinatorics I+II.

On a Personal Level ...

Personal Philosophy:



Modern Science:

- Specialized.
- Compartmentalized.
- Competitive.
- "Megaconferences".

(I hope that I convinced you that) sometimes it pays to straddle multiple scientific topics!

Hvala/Thank You !

References (for own work)

1. G. Istrate, A. Crãciun. Proof Complexity and the Kneser-Lovász Theorem. SAT'2014.
2. J. Aisenberg, M.L. Bonet, S. Buss, A. Crãciun and G. Istrate. Short Proofs for Kneser-Lovász formulae. ICALP'2015, journal version in Information and Computation 2018.
3. G. Istrate, C. Bonchiş, A. Crãciun. Kernelization, Proof Complexity and Social Choice. ICALP'2021. Journal version in progress.
4. G. Istrate (manuscript in progress).