From simple combinatorial statements with difficult mathematical proofs to hard instances of SAT

Gabriel Istrate

West University of Timişoara, Romania

Universitatea de Vest din Timişoara

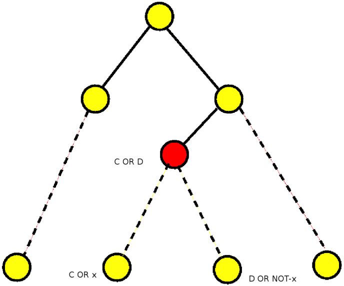
(joint work with Adrian Crãciun)

Caution

- This talk: "Theory A" (proof complexity), unpublished work.
- Naturally continues with experimental work on SAT benchmarks.
- One-line soundbite: Do combinatorial statements with difficult (mathematical) proofs correspond to "hard" instances of SAT ?
- I am not solving any major open problem in computational complexity

Reminder: Propositional proof complexity

- Proving that a formula is not satisfiable seems "harder" than finding a solution.
- Possible: proof systems for unsatisfiability, e.g. resolution
- $C \vee x, D \vee \overline{x} \rightarrow (C \vee D), x, \overline{x} \rightarrow \square$.
- Complexity= minimum length of a resolution proof.
- Lower bound for the running time of all DPLL algorithms !



Reminder: Propositional proof complexity (II)

- Resolution proof size may be exponential
- E.g. Pigeonhole formula(s): PHP_n^{n-1} (Haken)
- $X_{i,j} = 1$ "pigeon i goes to hole j".
- $X_{i,1} \vee X_{i,2} \vee \ldots \vee X_{i,n-1}$, $1 \leq i \leq n$ (each pigeon goes to (at least) one hole)
- $\overline{X_{k,j}} \vee \overline{X_{l,j}}$ (pigeons k and l do not go together to hole j).
- Resolution: clausal formulas. Stronger proof systems ?

Boundaries of proof complexity: Frege proofs

- Example, for concreteness [Hilbert Ackermann]
  - propositional variables $p_1, p_2, \ldots$.
  - Connectives $\neg, \vee$.
  - Axiom schemas:
    1. $\neg(A \vee A) \vee A$
    2. $\neg A \vee (A \vee B)$
    3. $\neg(A \vee B) \vee (B \vee A)$
    4. $\neg(\neg A \vee B) \vee (\neg(C \vee A) \vee (C \vee B))$
  - Rule: From A and $\neg A \vee B$ derive B.
- Cook-Reckhow: all Frege proof systems equivalent (polynomially simulate each other)
- Can prove *PHP* in polynomial size (Buss).
- Still exponential l.b. ($2^{n^{\epsilon}}$) if we restrict formula depth (bounded-depth Frege)

Boundary of knowledge: Frege proofs (II)

- PHP (Buss): proof by counting
- Usual proof by induction: exponential size in Frege: reduction causes formula size to increase by a constant factor at every reduction step.
- Polynomial if we allow introducing new variables: $X \equiv \Phi(\overline{Y})$.
- Frege + new vars: extended Frege

Our original idea/motivation

- Open question: Is extended Frege more powerful than Frege ?
- Most natural candidates for separation turned out to have subexponential Frege proofs.
- Perhaps translating into SAT a mathematical statement that is (mathematically) hard to prove would yield a natural candidate for the separation.
- Didn't quite work out: Our examples probably harder than extended Frege.

Kneser's Conjecture

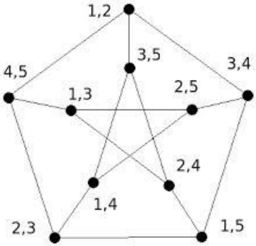
- Stated in 1955 (Martin Kneser, Jaresbericht DMV)
- Let $n \geq 2k - 1 \geq 1$. Let $c : \binom{n}{k} \to [n - 2k + 1]$. Then there exist two disjoint sets A and B with $c(A) = c(B)$.

Kneser's Conjecture

- Stated in 1955 (Martin Kneser, Jaresbericht DMV)
- 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)
- $k \geq 4$ only proved in 1977 (Lovász) using Algebraic Topology.
- Combinatorial proofs known (Matousek, Ziegler). "hide" Alg. Topology
- No "purely combinatorial" proof 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.

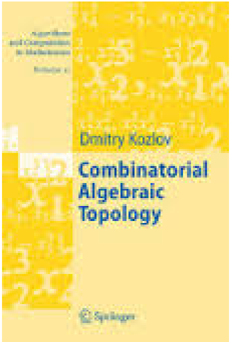


Stronger form: Schrijver's Theorem

- inner cycle in Petersen's graph already chromatic number three.
- $A \in \binom{n}{k}$ stable if it doesn't contain consecutive elements i, $i+1$ (including $n, 1$).
- Schrijver's Theorem: Kneser's conjecture holds when restricted to stable sets only.

Algebraic topology and graph colorings

- Dolnikov's theorem: generalization, lower bounds on the chromatic number of an arbitrary graph.
- In general not tight.
- Many other extensions.



Lovász-Kneser's Thm. as an (unsatisfiable) propositional formula

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

Our results in a nutshell

- $Kneser_{k,n}$ reduces to (is a special case of) $Kneser_{k+1,n-2}$.
- Thus all known lower bounds that hold for PHP (resolution, bd. Frege) hold for any $Kneser_k$.
- Cases with combinatorial proofs:
  - $k = 2$: polynomial size Frege proofs
  - $k = 3$: polynomial size extended Frege proofs
- $k \geq 4$: polynomial size implicit_2 extended Frege proofs
- Implicit proofs: Krajicek (2002). Very powerful proof system(s). AFAIK: first concrete example.

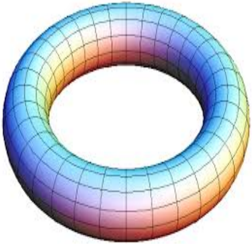
Significance

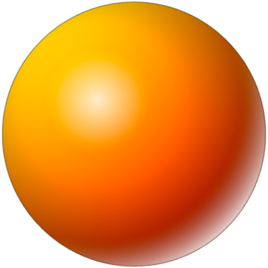
- Proof complexity: counterpart, expressibility in (versions of) bounded arithmetic
- Reverse mathematics: what is the weakest proof system that can prove a certain result ?
- Stephen Cook: "bounded reverse mathematics"

- Implicit proofs seem to be needed for simulating arguments involving algebraic topology.
- Reasons: ~~exponentially large objects~~ and nonconstructive methods
- CONJECTURE: For $k \geq 4$ $Kneser_{k,n}$ requires exponential-size (extended) Frege proofs

What is Algebraic Topology and why can it prove lower bounds on chromatic numbers ?

- Two objects similar if can continuously morph one into the other
- Cannot turn a donut into a sphere: Hole is an "obstruction" to contracting a circle going around the torus to a point.
- Can do that on a sphere.
- Continuous morphing should preserve contractibility.



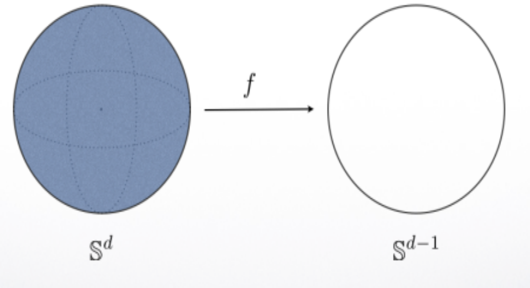


How do we "measure" the "number of holes" (and other properties) ?

- algebraic objects (groups)
- Functorial: $G \to H$ implies $F(G) \to F(H)$.
- If $K \to F(G)$ but $K \not\to F(H)$ then K acts as an obstruction to $G \to H$
- Coloring = morphism of graphs.

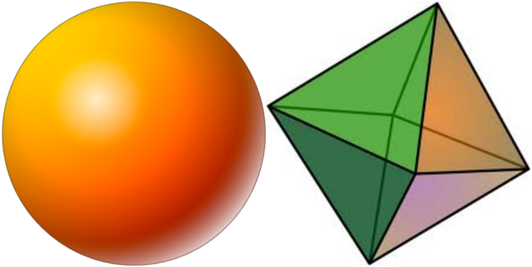
Ingredient of Kneser proof: Borsuk-Ulam Thm.

- Cannot map continuously and antipodally n-dim. sphere into a sphere of lower dimension (or ball into sphere)
- Obstruction: largest dimension of sphere that can be embedded continuously and antipodally into $F(G)$. As long as $F(K_m)$ "is a sphere".



From continuous to discrete

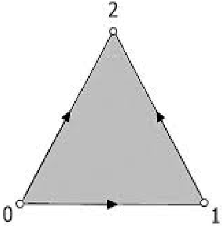
- A sphere is topologically equivalent to an octahedron
- simplicial complex: every subset of a face is a face.
- Simplex: purely combinatorially (sets that are simplices)

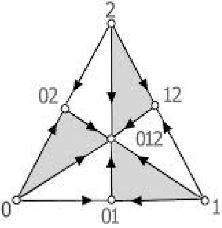


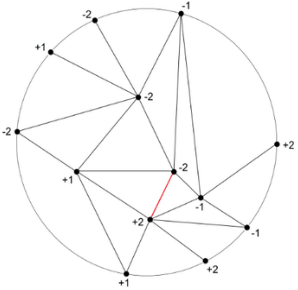
- Vertices: $\{\pm 1, \pm 2, \ldots, \pm n\}$.
- Faces: subsets that do not contain no i and $-i$.
- Exponentially (in n) many faces !

Discrete Borsuk-Ulam: 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.







KNESER FROM TUCKER ($k \geq 4$)

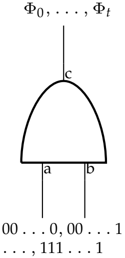
- Simulate "combinatorial" proof of Kneser (combination of two mathematical proofs)
- Tucker's lemma: unsatisfiable propositional formula. $Kneser_{k,n}$: variable substitution.
- barycentric dimension ⇒ exponentially large formula !
- Kneser follows from a new "low dimensional" Tucker lemma.
- Avoid barycentric subdivision. Instead (k+k) "skeleton"

Kneser from Tucker ($k \geq 4$)

- Second obstacle: Tucker lemma is nonconstructive (PPAD complete).
- Given an (exponential size) graph with one vertex of odd degree, find another node of odd degree
- For Kneser: this exponential graph has very regular structure.

Implicit proofs

- Krajicek (J. Symb. Logic 2004).
- Hierarchy: $iEF, i_2EF, i_3EF, \ldots$.
- ridiculously powerful: implicit resolution $\equiv$ extended Frege.
- poly-size boolean circuit that is generating all formulas in an extended Frege proof + correctness proof
- if correctness proof itself implicit $\Rightarrow$ second level. Correctness proof second level $\Rightarrow$ third level . . .



Implicit proofs: Kneser

- polynomial number of output gates $\Rightarrow \Phi_0, \ldots, \Phi_t$ "small"
- extended Frege: renaming keeps formulas small.
- implicit proofs allows us to generate a proof of the odd degree argument
- soundness: exponentially large (but regular) $\Rightarrow$ Kneser: second level

REDUCING $Kneser_{n,k+1}$ TO $Kneser_{n-2,k}$

- There exists a variable substitution $\Phi_k : Var(Kneser_{n,k+1}) \to Var(Kneser_{n-2,k})$ s.t. $\Phi_k(Kneser_{n,k+1})$ consists precisely of the clauses of $Kneser_{n-2,k}$ (perhaps repeated and in a different order)
- Let $A \in \binom{n}{k+1}$. Define $\Phi_k(X_{A,i})$ by:
  - **Case 1:** $A_{\leq k} \subseteq [n-2]$**:** $\Phi_k(X_{A,i}) = Y_{A_{\leq k},i}$
  - **Case 2:** $A_{\leq k} \not\subseteq [n-2]$**:** $(n-1, n \in A)$
    Let $A = P \cup \{n-1, n\}$, $|P| = k-1$. Let $\lambda = max\{j : j \leq n-2, j \notin P\}$. Define $\Phi_k(X_{A,i}) = Y_{P\cup\{\lambda\},i}$
- Clause $X_{A,1} \vee X_{A,2} \vee \ldots \vee X_{A,n-2k+1}$ maps to $Y_{B,1} \vee Y_{B,2} \vee \ldots \vee Y_{B,n-2k+1}$, $B = A$ (Case 1).
- Clauses $\overline{X_{A,i}} \vee \overline{X_{B,i}}$ $(A \cap B = \emptyset)$ map to $\overline{Y_{C,i}} \vee \overline{Y_{D,i}}$
- Case 2 cannot happen for both A and B. By case analysis $C \cap D = \emptyset$.

Comments on (other) proofs

- Lower bounds Schrijver: Same substitution, slightly more complicated argument.
- $k = 2$: counting proof, Stahl+ Buss PHP.
- For any color class $c^{-1}(\lambda)$ one of the following is true (assuming conclusion of Kneser does not hold):
  - $|c^{-1}(\lambda)| \leq 3$.
  - All sets $B \in c^{-1}(\lambda)$, $|c^{-1}(\lambda)| \geq 4$, have one element in common (call such an element special).
  - Frege systems can "count" (employing techniques developed by Buss) the number of special elements.
- $k = 3$: Counting approach fails (technical reasons), have to settle for extended Frege.

From Kneser-like results to hard SAT instances ?

- $2^{\Omega(n)}$ resolution complexity. Are they hard in practice ?
- At this point: only idea for subsequent work
- Want: small formulas.
- $Kneser_{n,k}$: $\sim n^{k+1}$ variables, even more clauses.
- Schrijver ? Other versions of Dolnikov's Theorem ? expander graph with tight bounds on the chromatic number
- Better encodings ? All intuitions should apply.
- Kneser, stable Kneser graphs: symmetries well understood. But: reason for unsatisfiability is more global

Further possible work

- Other proof systems: e.g. cutting planes (k=2), polynomial calculus, etc.
- (in progress) Topological obstructions: from graph coloring to CSP.
- Logics for implicit proof systems ?
- Topological arguments as sound (but incomplete) implicit proof systems
  - if $K \not\to L$ then a "proof of $A \not\to B$" is a pair of embeddings $(K \to A), (B \to L)$.
  - Checking soundness ($K \not\to L$) may not be polynomial. If K, L "standard objects" we could omit proof of $K \not\to L$ from complexity
- Automated theorem proving ?

Thank you. Questions ?