The complexity of promise SAT on non-Boolean domains

Alex Brandts, Marcin Wrochna, and Standa Živný

2-SAT vs 3-SAT

- 2-SAT is in P, 3-SAT is NP-hard – what happens in between?

For a k-CNF formula, $(1, g, k)$-SAT asks to distinguish between the two cases:

  - There exists a truth assignment satisfying at least g literals in every clause
  - The formula is unsatisfiable

Theorem: [Austrin Guruswami Håstad] When $\frac{g}{k} \geq \frac{1}{2}$, $(1, g, k)$-SAT is in P and otherwise is NP-hard

- How to extend this problem to larger domains?

$(1, g, k)$-SetSAT

- Domain $[d]$
- Literals $S_a(x)$ satisfied iff $x \neq a$
- Clauses $(S_{a_1}(x_1), S_{a_2}(x_x), \ldots, S_{a_k}(x_k))$
- $(1, g, k)$-SetSAT asks to distinguish the two cases:
  - there exists an assignment to the variables such that at least g literals in each clause are satisfied
  - the formula is unsatisfiable

Main result: $(1, g, k)$-SetSAT with set size s and domain size $s+1$ is solvable in polynomial time if $\frac{g}{k} \geq \frac{s}{s+1}$ and is NP-hard otherwise.

- With $s = 1$ we recover the results for Boolean $(1, g, k)$-SAT

Tractability

Algorithm 1 Randomised algorithm for $(1, g, k)$-SetSAT with $\frac{g}{k} \geq \frac{s}{s+1}$.

$x \leftarrow$ arbitrary assignment
while x does not satisfy input formula ϕ **do**
    Arbitrarily pick a falsified clause C
    Randomly choose from C a literal $S(x_i)$
    Randomly choose a value for x_i so that $S(x_i)$ is satisfied
return x

Analysis

- Suppose ϕ has a g satisfying assignment x^*.
- Let x^t be assignment at iteration t
- As t increases, the distance from x^t to x^* decreases in expectation
- Biased random walk reaches x^* with constant probability in $O(n^2)$ steps

Tractability via polymorphisms

- What polymorphisms does $(1, g, k)$-SetSAT have when $\frac{g}{k} \geq \frac{s}{s+1}$?
- All polymorphisms of SetSAT are conservative
- Plurality functions are polymorhpisms

$$f(x_1, \ldots, x_m) = \text{argmax}_{a \in [d]} \{\# \text{ of occurrences of } a \text{ in } (x_1, \ldots, x_m)\}$$

- When $\frac{g}{k} > \frac{s}{s+1}$, we have plurality pols of all arities $\Rightarrow$ solvable by BLP
- When $\frac{g}{k} = \frac{s}{s+1}$, we have plurality of arity $\not\equiv 0 \pmod{s+1}$ and no symmetric pols of arity divisible by $s+1$ $\Rightarrow$ solvable by BLP+AIP

Hardness

- What polymorphisms remain when $\frac{g}{k} < \frac{s}{s+1}$?

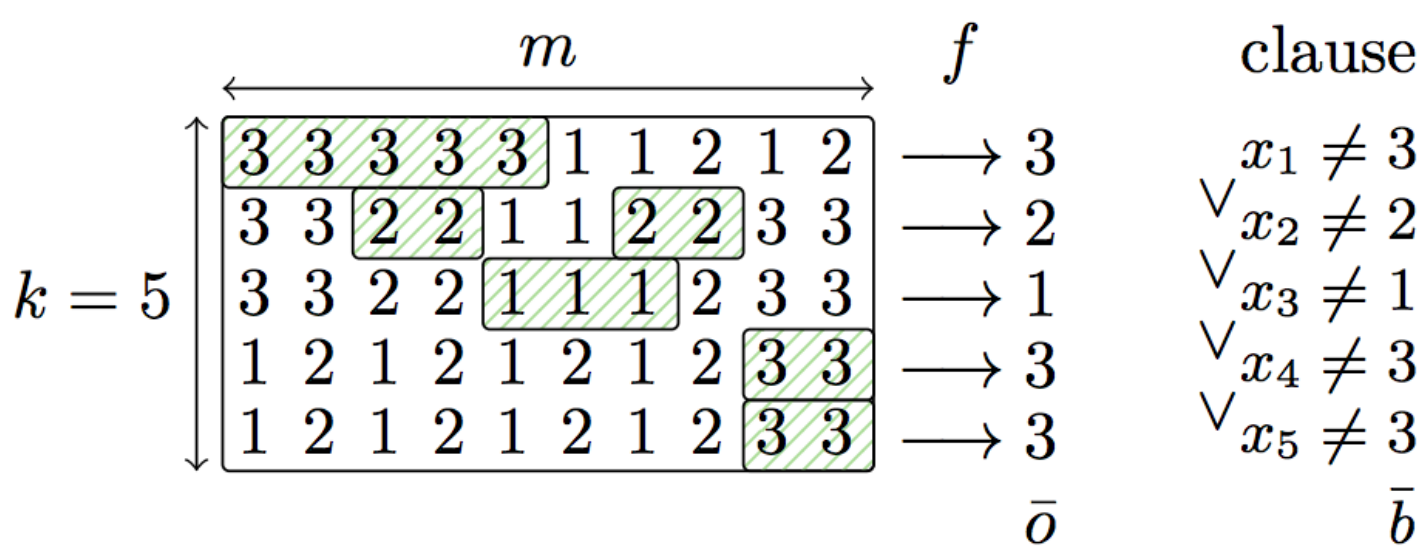
- All plurality pols now have bounded essential arity, but there still exist pols of unbounded essential arity

- Pols don't satisfy sufficient hardness conditions involving fixing sets, avoiding sets, ϵ-robustness, lack of Olšák polymorphisms

- We need a new hardness source, or a new property of pols to exploit

Smug sets

- S is a *smug set* if there is an input vector v such that $S = \{i | v_i = f(v)\}$



Proposition: A function $f : [s+1]^m \to [s+1]$ is a polymorphism of $(1, g, k)$-SetSAT iff there is no multiset $\{S_1, \ldots, S_k\}$ of smug sets of f, such that each coordinate $\ell \in [m]$ is contained in at most $k - g$ of them.

Hardness source: layered label cover

- ℓ layers of variables $X_0, \ldots, X_\ell$ with range $[m]$
- Constraints are functions from $x \in X_i$ to $y \in X_j$, $i < j$
- Constraint $\phi_{x \to y}$ is satisfied by assignment σ if $\sigma(y) = \phi_{x \to y}(\sigma(x))$
- A chain is a sequence of variables $x_i \in X_i$ for $i = 0, \ldots, \ell$ with constraints $\phi_{x_i \to x_j}$ between them, for $i < j$
- A chain is weakly satisfied if at least one of these constraints is satisfied

Theorem: For every ℓ and $\epsilon > 0$, there is an m such that it is NP-hard to distinguish ℓ-layered label cover instances with domain size m that are fully satisfiable from those where not even an ϵ-fraction of all chains is weakly satisfied.

Smug sets and hardness

Corollary: Suppose there are constants k, ℓ such that the following holds, for every $f \in \mathrm{Pol}(\mathbf{A}, \mathbf{B})$:

- f has a smug set of at most k coordinates
- f has no family of more than ℓ pairwise disjoint smug sets
- if $g \xrightarrow{\pi} f$ and S is a smug set of g, π^{-1}(S) is a smug set of f

Then $\mathrm{PCSP}(\mathbf{A}, \mathbf{B})$ is NP-hard, for large enough m.

- Exact definition of *smug* is irrelevant, as long as it satisfies these three conditions

Bounded number of disjoint smug sets

Proposition: For every polymorphism f of $(1, g, k)$-SetSAT with domain size $s+1$, if $S_1, \ldots, S_t$ are disjoint smug sets of f, then $t < \frac{k}{k-g}$.

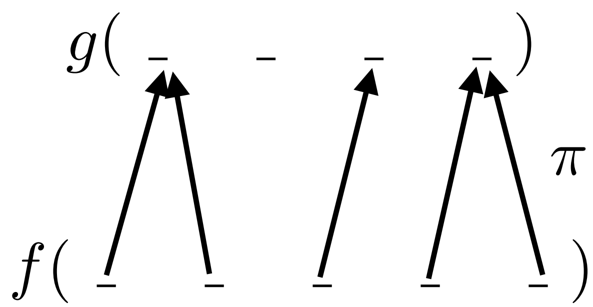
Proof:

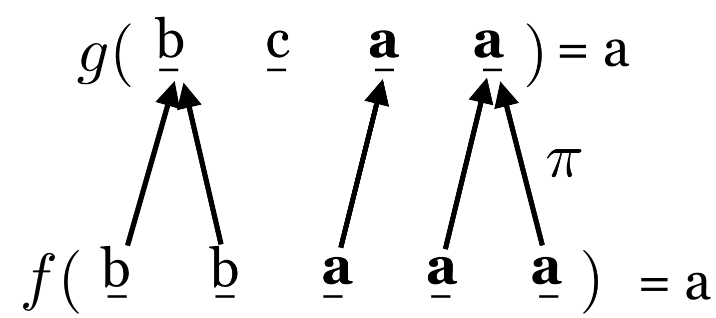
- Suppose $t \geq \frac{k}{k-g}$
- Build a multiset containing each S_i up to $k-g$ times until we have exactly k in total
- This gives a multiset of k smug sets such that every coordinate is contained in at most $k-g$ of them
- Contradicts earlier proposition:

  Proposition: A function $f : [s+1]^m \to [s+1]$ is a polymorphism of $(1, g, k)$-SetSAT iff there is no multiset $\{S_1, \ldots, S_k\}$ of smug sets of f, such that each coordinate $\ell \in [m]$ is contained in at most $k-g$ of them.

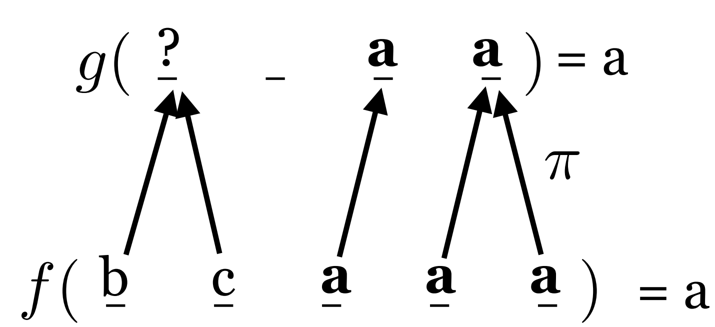
Smug sets preserved by minor preimages

g is a minor of f: $g(x_1, \ldots, x_m) \approx f(x_{\pi(1)}, \ldots, x_{\pi(n)})$





Not preserved by minor images:



Existence of small smug sets

Proposition: Let $\frac{g}{k} < \frac{s}{s+1}$. Every polymorphism of $(1, g, k)$-SetSAT with domain size $s+1$ has a smug set of size at most g.

Proof:

- Construct small set from collection of minimal smug sets, using conservativity and minimality to enforce outputs of pols

- Much more involved than proving the other two properties

Generalizations and open problems

- What if the literals are drawn from an arbitrary family of sets $\mathcal{L} \subseteq \mathcal{P}([d])$?

- Some reductions and easy cases, eg if $\bigcap_{L \in \mathcal{L}} L \neq \varnothing$ then problem is tractable

Conjecture: Let $\mathcal{L} \subseteq \mathcal{P}([d])$ and let $s_{\max}$ be the size of the largest set in $\mathcal{L}$. If $\bigcap_{L \in \mathcal{L}} L = \varnothing$ then $(1, g, k, \mathcal{L})$-SetSAT is tractable iff $\frac{g}{k} \geq \frac{s_{\max}}{s_{\max}+1}$.

- Tractability: randomized algorithm still works for $\frac{g}{k} \geq \frac{s_{\max}}{s_{\max}+1}$

- Hardness: polymorphisms are no longer conservative, so smug sets can't immediately be used

Hypergraph colouring

Theorem: [Austrin Guruswami Håstad] For $g \geq 1$, given a $(2g+1)$-uniform hypergraph that admits a 2-colouring of discrepancy 1 (smallest possible), it is NP-hard to find a non-monochromatic 2-colouring.

Conjecture: Given an $(s+1)r + a$ uniform hypergraph H with $1 \leq a \leq s$, it is NP-hard to distinguish the two cases:

- There exists an $s+1$ colouring of H of discrepancy at most 1
- Every $s+1$ colouring of H creates a monochromatic hyperedge

- If we are instead promised discrepancy 0, the problem is tractable by a reduction to SetSAT
- Given an $(s+1)r$ uniform hypergraph H, the two cases can be distinguished in polynomial time:
    - There exists an $s+1$ colouring of H of discrepancy 0
    - Every $s+1$ colouring of H creates a monochromatic hyperedge

Hypergraph colouring

Conjecture: Given an $(s+1)r+a$ uniform hypergraph H with $1 \leq a \leq s$, it is NP-hard to distinguish the two cases:

- There exists an $s+1$ colouring of H of discrepancy at most 1
- Every $s+1$ colouring of H creates a monochromatic hyperedge

- This conjecture would imply all the SetSAT hardness results

- Challenges: polymorphisms not conservative, PCSP is much more symmetric

Thank you!