Cyclic operations in promise constraint satisfaction problems

Libor Barto

Department of Algebra, Charles University, Prague

Dagstuhl CSP seminar, 3–8 June 2018

CoCoSym: Symmetry in Computational Complexity

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 771005)

- for finite relational structure $\mathbb{A}$
  - $\mathrm{CSP}(\mathbb{A})$: given $\mathbb{X}$ find $\mathbb{X} \to \mathbb{A}$
  - …a computational problem, one for each $\mathbb{A}$
  - **Example:** Find a 3-coloring of a graph (for $\mathbb{A} = \mathbb{K}_3$)
  - $\mathrm{Pol}(\mathbb{A}) = \{f : \mathbb{A}^n \to \mathbb{A}\}$ polymorphisms
  - **Fact:** it is closed under composition (it is a clone)
- complexity of $\mathrm{CSP}(\mathbb{A})$ depends only on
  - $\mathrm{Pol}(\mathbb{A})$ [Jeavons'98]
  - identities in $\mathrm{Pol}(\mathbb{A})$ [Bulatov, Jeavons, Krokhin'05]
  - height one identities in $\mathrm{Pol}(\mathbb{A})$ [B, Opršal, Pinsker'17]
- $\mathrm{CSP}(\mathbb{A})$ is
  - **hard** if polymorphisms don't satisfy some "nontrivial" height one identities
  - **easy** if they do
  - here "nontrivial" means not satisfiable by projections (aka dictators) [Bulatov'17]; [Zhuk'17]

- identity is universally quantified equation
- (identification) minor of $f : A^n \to A$ is an operation $g : A^m \to A$ defined by

$$g(x_1, \dots, x_m) = f(\text{ variables })$$

  e.g. $g(x, y) = f(x, y, x, x, y)$

- height one identity is of the form

$$f(\text{ variables }) = g(\text{ variables })$$

- i.e. equality between identification minors of f and g
- **Note:** makes sense for $f, g : A^n \to B$

- for finite relational structures $\mathbb{A}, \mathbb{B}$ with $\mathbb{A} \to \mathbb{B}$
  - $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$: given $\mathbb{X}$ such that $\mathbb{X} \to \mathbb{A}$ find $\mathbb{X} \to \mathbb{B}$
  - ... a computational problem, one for each pair $\mathbb{A}, \mathbb{B}$
  - **Example:** Find a 4-coloring of a 3-colorable graph
  - $\mathsf{Pol}(\mathbb{A}, \mathbb{B}) = \{f : \mathbb{A}^n \to \mathbb{B}\}$ polymorphisms
  - **Observe:** general composition does not make sense
  - **Fact:** closed under identification minors
    (it is a clonoid (?), minion (?), proclone (?)...)
- complexity of $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$ depends only on
  - $\mathsf{Pol}(\mathbb{A}, \mathbb{B})$ [Brakensiek, Guruswami'16]
  - height one identities in $\mathsf{Pol}(\mathbb{A}, \mathbb{B})$ [Bulín, Krokhin, Opršal]
- $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$ is
  - **hard** if polymorphisms don't satisfy
    some "nontrivial" height one identities
  - **easy** if they do
  - here "nontrivial" means ???

Hard and easy CSPs

- **hardness:**
  no nontrivial height one identities + abstract nonsense
  $\Rightarrow$ reduction from any NP-hard CSP (e.g. Label Cover)

- **easiness:**
  nontrivial height one identities
  $\Rightarrow$ stronger identities (e.g. cyclic
  $f(x_1, \ldots, x_n) = f(x_2, \ldots, x_n))$
  $\Rightarrow$ algorithm

There is no gap between "nontrivial" in the two cases

- **hardness** no "nontrivial" height one identities + nonsense
  $\Rightarrow$ reduction from NP-hard Gap Label Cover problem

  Given a system of height one identities
  which are satisfiable by projections
  Find an assignment operations $\to$ projections
  which satisfies at least $1/100$ identities

- identities of "permutation type" seem especially important (see Unique Games)

- **easiness** "nontrivial" height one identities $\Rightarrow$ algorithm

There is a gap between "nontrivial" in the two cases

- **CSP easiness** nontrivial identities $\Rightarrow$ stronger identities $\Rightarrow$ algorithm
- **PCSP easiness** nontrivial identities $\Rightarrow$ algorithm

Contributions

- missing results: identities $\Rightarrow$ stronger identities

  Contribution: monotone Boolean cyclic $\Rightarrow$ threshold

- algorithms in PCSPs based on **infinite** domain CSPs (LP, Gauss over $\mathbb{Z}$)

  Contribution: PCSP(1-in-3, NAE) not solvable using finite domain CSP (in some sense)

Cyclic monotone Boolean operations

Boolean operation $f : \{0,1\}^n \to \{0,1\}$ is

- cyclic if $f(x_1, x_2, \dots, x_n) = f(x_2, \dots, x_n, x_1)$
- fully symmetric if $f(x_1, x_2, \dots, x_n) = f(x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(n)})$ for each $\pi \in S_n$
- threshold if it equals thr_α for some α where

$$\mathrm{thr}_\alpha(x_1, \dots, x_n) = 1 \text{ iff } \sum x_i > \alpha n$$

- monotone if it preserves $\leq$ where $0 \leq 1$

Note: threshold = monotone + fully symmetric

Theorem

For each k there exists l such that every cyclic monotone Boolean operation of arity $n \geq l$ has an identification minor of arity $\geq k$ which is a threshold operation.

∞-many threshold polymorphisms $\Rightarrow$ tractability of PCSP
[Brakensiek,Guruswami'16]

Corollary

Let $\mathbb{A} \to \mathbb{B}$ be Boolean, containing $\leq$. If $\mathrm{Pol}(\mathbb{A}, \mathbb{B})$ contains ∞-many cyclic operations, then $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$ is tractable.

How far from dichotomy for monotone Boolean PCPS?

Let $f : \{0,1\}^n \to \{0,1\}$ monotone and $p \in [0,1]$

- choose $x_1, \ldots, x_n \in \{0,1\}$ independently
  - $x_i = 1$ with probability p
  - $x_i = 0$ with probability $1 - p$
- $E_f(p) =$ expected value of $f(x_1, \ldots, x_n)$
- $I_f(p,i)$ influence of the i-th variable
  $=$ probability that $f(x_1, \ldots, x_n)$ changes when x_i is changed
- $I_f(p) := \sum_i I_f(p,i)$ total influence

Theorem ("Russo's Lemma")

$E_f'(p) = I_f(p)$

Theorem ("KKL Theorem" [Kahn, Kalai, Linial'88])

$\exists i \quad I_f(p,i) \geq C\ E_f(p)(1 - E_f(p))\ \frac{\log n}{n}$

Proving: Cyclic monotone $f : \{0,1\}^n \to \{0,1\}$ of sufficiently large arity n has a threshold minor of arity ≥ 10.

Russo's Lemma: $E_f'(p) = I_f(p)$
KKL Theorem: $\exists i \;\; I_f(p,i) \geq C \;\, E_f(p)(1 - E_f(p)) \;\, \log n/n$

- take p such that $E_f(p) = 0.5$, say $E_f(0.36) = 0.5$
- f cyclic so $I_f(p,i) = I_f(p,j)$ so $I_f(p) = nI_f(p,i)$
- Russo+KKL: $E_f'(p) = I_f(p) \geq CE_f(p)(1 - E_f(p)) \;\, \log(n)$
- if $0.00001 \leq E_f(p) \leq 0.99999$ then $E_f'(p) \geq D\log(n)$
- n large $\Rightarrow$
  - if $p < 0.35$ then $E_f(p) < 0.00001$
  - if $p > 0.37$ then $E_f(p) > 0.99999$

Proof 2/2

$p < 0.35 \Rightarrow E_f(p) < 0.00001 \qquad p > 0.37 \Rightarrow E_f(p) > 0.99999$

- choose a random 10-ary minor of f
  ie. define $g(x_1, \ldots, x_{10}) = f(y_1, \ldots, y_n)$ where
  y_i are chosen uniformly independently from $\{x_1, \ldots, x_{10}\}$
- **Aim**: $\mathrm{P}(g = \mathrm{thr}_{0.35}) > 0$
- $\mathrm{Exp}(g(1,1,1,0,0,0,\ldots,0)) = E_f(3/10) < 0.00001$
- $\mathrm{Exp}(g(1,1,1,1,0,0,\ldots,0)) = E_f(4/10) > 0.99999$
- Expected value of

  $$V := g(1,1,1,0,0,\ldots,0) + g(1,1,0,1,0,\ldots,0) + \cdots + (1 - g(1,1,1,1,0,\ldots,0)) + (1 - g(1,1,1,0,1,0,\ldots,0)) + \ldots$$

  is at most $\binom{10}{3}0.00001 + \binom{10}{4}0.00001 < 1$
- So $\mathrm{P}(V = 0) > 0$
- But $\mathrm{P}(V = 0) = \mathrm{P}(g = \mathrm{thr}_{0.35})$

Finite domain CSP is insufficient for PCSP

$\mathrm{PCSP}(\mathbb{A},\mathbb{B})$ tractable (e.g. ∞-many threshold polymorphisms)

- $\mathbb{A} = (\{0,1\};\ \{(1,0,0),(0,1,0),(0,0,1)\})$ (1-in-3)
- $\mathbb{B} = (\{0,1\};\ \{0,1\}^3 \setminus \{(0,0,0),(1,1,1)\}$ (NAE)

Theorem

There is no tractable finite–domain $\mathrm{CSP}(\mathbb{C})$ such that height one identities satisfied in Pol($\mathbb{C}$) *are satisfiable in* Pol($\mathbb{A},\mathbb{B}$).

Proof:

- Relational counterpart of height one identities:
  [Bulín, Krokhin, Opršal]

  pp-interpretation + generalization of homomorphic equivalence
  (in [Brakensiek,Guruswami'08] called "promise embedding")
- cyclic operations + work

Questions

- what is "nontrivial height one identities"?
- are existing guesses sufficient, at least for
  - monotone Boolean PCSPs?
  - Boolean PCSPs?
  - $\mathrm{PCSP}(\mathbb{K}_n, \mathbb{K}_m)$

Questions

- what is "nontrivial height one identities"?
- are existing guesses sufficient, at least for
  - monotone Boolean PCSPs?
  - Boolean PCSPs?
  - $\mathrm{PCSP}(\mathbb{K}_n, \mathbb{K}_m)$

Thank you!