

Generic Expression Hardness Results for Primitive Positive Formula Comparison

Hubie Chen

Universitat Pompeu Fabra
Barcelona, Spain

August 2011 - Fields Institute

Joint work with Simone Bova (Vanderbilt Univ., US)
and Matt Valeriote (McMaster Univ., Canada)

Primitive Positive Formulas

Definition

A *primitive positive formula* is a first-order formula defined from atomic formulas, equality of vars., conjunction (\wedge), and existential quantification (\exists).

In prenex form:

$$\exists v_1 \dots \exists v_n (\alpha_1 \wedge \dots \wedge \alpha_m)$$

where each α_i is atomic ($R(w_1, \dots, w_k)$ or an equality ($v = w$))

- Includes *conjunctive queries* from relational database theory
- Have been studied complexity-theoretically from a number of perspectives
 - Deciding if a primitive positive sentence holds on a structure is a formulation of the *constraint satisfaction problem*

Problems Studied

We study the complexity of two basic problems.

Problems studied

Equivalence: given two pp formulas $\phi(X), \phi'(X)$ and a structure \mathbf{B} , are the formulas equivalent over \mathbf{B} ?

(does it hold, for all $f : X \rightarrow B$, that $\mathbf{B}, f \models \phi$ iff $\mathbf{B}, f \models \phi'$?)

Containment: given two pp formulas $\phi(X), \phi'(X)$ and a structure \mathbf{B} , is ϕ contained in ϕ' , over \mathbf{B} ?

(does it hold, for all $f : X \rightarrow B$, that $\mathbf{B}, f \models \phi$ implies $\mathbf{B}, f \models \phi'$?)

Problems Studied

We study these problems with respect to fixed structures. Let \mathbf{B} be a structure.

Parameterized versions

PPEQ(\mathbf{B}): Equivalence problem over \mathbf{B} .

(Given two formulas $\phi(X), \phi'(X)$, decide if ϕ, ϕ' equal over \mathbf{B} .)

PPCON(\mathbf{B}): Containment problem over \mathbf{B} .

(Given two formulas $\phi(X), \phi'(X)$, decide if ϕ contained in ϕ' over \mathbf{B} .)

- We have two *families* of problems
- For *each* \mathbf{B} , we can talk about complexity of PPEQ(\mathbf{B}), PPCON(\mathbf{B})
- We want to understand interplay between
structural properties of $\mathbf{B} \leftrightarrow$
complexity of PPEQ(\mathbf{B}), PPCON(\mathbf{B})

Observations

- We focus on relational structures \mathbf{B} that are finite in two senses:
 - the universe B is finite
 - the signature on which \mathbf{B} is defined is finite
- Obs: For any such \mathbf{B} , the problems $\text{PPEQ}(\mathbf{B})$, $\text{PPCON}(\mathbf{B})$ are in Π_2^P

Why: Let $\phi(X), \phi'(X)$ be pp formulas with quantified vars. V, V' . To check if $\phi(X)$ contained in $\phi'(X)$, check:

for all assignments $f : X \rightarrow B$ and $g : V \rightarrow B$,
 f, g satisfies $\text{quant-free}(\phi) \rightarrow$ **exists** g' sat. $\text{quant-free}(\phi')$

- Obs: For each struct. \mathbf{B} , it holds that $\text{PPCON}(\mathbf{B})$ polytime reduces to $\text{PPEQ}(\mathbf{B})$

Proof: Map an instance (ϕ, ϕ') of $\text{PPCON}(\mathbf{B})$ to the instance $(\phi, \phi \wedge \phi')$ of $\text{PPEQ}(\mathbf{B})$.

Summary of Results

- We present two general hardness results which have the form:
 - If \mathbf{B} satisfies condition (X1), then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are Π_2^P -hard (and hence Π_2^P -complete)
 - If \mathbf{B} satisfies condition (X2), then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are coNP-hard
- These results are optimal under two conjectures:
 - Suppose conjecture (C1). If \mathbf{B} does not satisfy condition (X1), then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are in coNP
 \Rightarrow get coNP/ Π_2^P -complete dichotomy
 - Suppose conjecture (C2). If \mathbf{B} does not satisfy condition (X2), then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are in P
 \Rightarrow get P/coNP-hard dichotomy
- Under the two conjectures, our hardness results thus imply a P/coNP-complete/ Π_2^P -complete trichotomy in the complexity of the problems
 - Full complexity classification on finite structures

Universal Algebra

- Our study makes use of universal-algebraic tools and notions
- An operation $f : B^m \rightarrow B$ is a *polymorphism* of a relation $U \subseteq B^k$ if for any choice of m tuples from U , applying f coordinatewise yields a tuple (s_1, \dots, s_k) also in U .

$$(t_{11}, t_{12}, \dots, t_{1k}) \in U$$

$$\vdots \quad \vdots \quad \ddots \quad \vdots$$

$$(t_{m1}, t_{m2}, \dots, t_{mk}) \in U$$

$$f \downarrow \quad f \downarrow \quad \dots \quad f \downarrow$$

$$(s_1, s_2, \dots, s_k) \in U$$

- An op. is a polymorphism of a structure if it is a polymorphism of all of its relations
Equivalently, $h : B^k \rightarrow B$ is a polymorphism of \mathbf{B} if it is a homomorphism $\mathbf{B}^k \rightarrow \mathbf{B}$

Universal Algebra

- Let $\text{Pol}(\mathbf{B})$ denote the set of polymorphisms of \mathbf{B} .
Let $\mathbb{A}_{\mathbf{B}}$ denote the algebra $(B, \text{Pol}(\mathbf{B}))$.
- Thm (follows from Geiger/Bodcharnuk et al., late 60s):
If for two structures \mathbf{B}, \mathbf{B}' it holds that $\mathbb{A}_{\mathbf{B}} = \mathbb{A}_{\mathbf{B}'}$, then
 - \mathbf{B} and \mathbf{B}' have the same pp-definable relations
 - $\text{PPEQ}(\mathbf{B}), \text{PPEQ}(\mathbf{B}')$ are interreducible; similarly, $\text{PPCON}(\mathbf{B}), \text{PPCON}(\mathbf{B}')$ interreducible
(wrt many-one polytime reduction)
- Recall: variety $\mathcal{V}(\mathbb{A})$ of an algebra \mathbb{A} is class of algebras derivable by taking homomorphic images, subalgebras, and products
- Thm: Let \mathbf{B} be a struct, and suppose $\mathbb{C} \in \mathcal{V}(\mathbb{A}_{\mathbf{B}})$.
If all \mathbb{C} -ops. are polymorphisms of struct \mathbf{D} ,
then $\text{PPEQ}(\mathbf{D})$ reduces to $\text{PPEQ}(\mathbf{B})$
(and same for $\text{PPCON}(\cdot)$)

- Def: *Congruence* of an algebra \mathbb{A} is an equivalence relation preserved by \mathbb{A} -ops (having all \mathbb{A} -ops as polymorphisms)
 - Equivalently: an equivalence relation of form
$$\{(a, a') \mid h(a) = h(a')\}$$
where h a homomorphism from \mathbb{A} to another algebra
- Congruences of an algebra form a lattice

First hardness result

- Tame congruence theory: associates a *typeset* to each finite algebra (based on congruence lattice) containing one or more of five *types*
- By extension, assigns a typeset to each variety V : union of all typesets of finite algebras in V
- Theorem: Let \mathbf{B} be a finite struct. If $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ admits the unary type, then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are Π_2^p -hard.

First hardness result

- Theorem: Let \mathbf{B} be a finite struct. If $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ admits the unary type, then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are Π_2^P -hard.
- $\text{CSP}(\mathbf{B})$: decide, for a given pp sentence ψ , if $\mathbf{B} \models \psi$
- Let \mathbf{B}^* be expansion of \mathbf{B} containing all constants
- G-set conjecture (Bulatov, Jeavons, Krokhin): if \mathbf{B}^* omits the unary type, $\text{CSP}(\mathbf{B}^*)$ is in P.
(otherwise, $\text{CSP}(\mathbf{B}^*)$ known to be NP-complete)
- Under the G-set conjecture, theorem yields Π_2^P -hard/coNP dichotomy:
 - If $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ omits the unary type, then $\mathcal{V}(\mathbb{A}_{\mathbf{B}^*})$ also does
 - By conjecture, $\text{CSP}(\mathbf{B}^*)$ is in P
 - $\text{PPEQ}(\mathbf{B})$ is in coNP: formulas $\phi(X), \phi'(X)$ are equiv. when
for all $f : X \rightarrow B, (\mathbf{B} \models \phi(f) \leftrightarrow \mathbf{B} \models \phi'(f))$
 - Conditions $\mathbf{B} \models \phi(f), \mathbf{B} \models \phi'(f)$ can be viewed as instances of $\text{CSP}(\mathbf{B}^*)$

What's going on here?

How does our hardness condition compare to the known hardness condition for the CSP?

- Reminder: \mathbf{B}^* is the expansion of \mathbf{B} by all constants
- Our thm: If var. of \mathbf{B} admits the unary type, then $\text{PPEQ}(\mathbf{B}), \text{PPCON}(\mathbf{B})$ are Π_2^P -hard
- Known from (Bulatov, Jeavons, Krokhin):
 - If var. of \mathbf{B}^* admits the unary type, then $\text{CSP}(\mathbf{B}^*)$ NP-hard
 - This hardness result can be applied to general structures via a transformation $\mathbf{A} \rightsquigarrow \mathbf{B}^*$ that preserves CSP complexity (i.e., $\text{CSP}(\mathbf{A}), \text{CSP}(\mathbf{B}^*)$ interreducible)
 - Transformation may reduce the universe of the structure
- **The two hardness conditions are different!**
 - Example: let \mathbf{A} be the structure on $\{0, 1\}$ containing all 3-ary relations R with $(0, 0, 0), (1, 1, 1) \in R$
 - Our hardness result applies to \mathbf{A}
 - On the other hand, $\text{CSP}(\mathbf{A})$ easy (instances always sat)

What's going on here?



- On instances of $\text{CSP}(\mathbf{A})$, one can “localize” the CSP to a subset of the universe
(general phenomenon: can pass to the core)
- This localization preserves the property of a pp-formula having a satisfying assignment
- ...but does not preserve the space of satisfying assignments
Problems studied here concern entire solution space!
- Open issue: do our hardness results have consequences for *other* problems involving the *global solution space* of CSPs?
- Remark: localization also fails for counting CSP –
another example of problem dealing with global solution space



- If you believe...
 - CSP admits dichotomy (on finite structures), and
 - $NP \neq coNP$
- AND you can prove $PPEQ(\mathbf{B})$ in $coNP$ for all \mathbf{B} not covered by our hardness result...
- Then you have proved the G-set conjecture!
- Proof: exercise

Second hardness result

- An algebra is *congruence modular* if its lattice of congruences satisfies modularity:

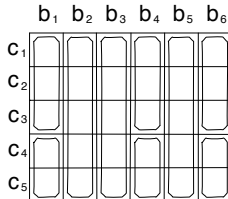
$$x \leq y \rightarrow x \vee (y \wedge z) = y \wedge (x \vee z)$$

- A variety is congruence modular if all of its members are congruence modular
- Theorem: Let \mathbf{B} be a finite struct. If $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ is not congruence modular, then $\text{PPEQ}(\mathbf{B})$, $\text{PPCON}(\mathbf{B})$ are coNP-hard.
- Theorem yields P/coNP-hard dichotomy under...

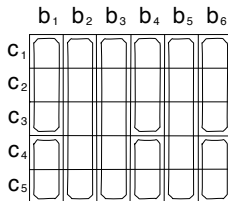
Second hardness result



- Theorem yields P/coNP-hard dichotomy under...
- Valeriote's Edinburgh conjecture: assuming \mathbf{B} has finite signature, if $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ is congruence modular, then $\mathbb{A}_{\mathbf{B}}$ has few subpowers
- Few subpowers: $\log \#S(\mathbb{A}_{\mathbf{B}}^n)$ grows polynomially in n , where $\#S(\mathbb{A})$ is the number of subalgebras of alg. \mathbb{A}
- Few subpowers is known to imply polytime decidability of PPEQ(\mathbf{B}), PPCON(\mathbf{B})



- A *pentagon* is a structure $\mathcal{P} = (P, \alpha, \beta, \gamma)$ such that α, β, γ are equivalence relations on P , and:
 - 1 $\alpha \leq \beta$
 - 2 $\beta \wedge \gamma = 0$
 - 3 $\beta \cdot \gamma = 1$
 - 4 $\alpha \vee \gamma = 1$
- A pentagon can be decomposed as $P = B \times C$ where β, γ are the kernels of the projections onto B, C
- Example: $B = \{b_1, \dots, b_6\}$, $C = \{c_1, \dots, c_5\}$.
Will sketch proof of coNP-hardness on this example pentagon.
- Columns $\equiv \beta$ -equiv. classes, rows $\equiv \gamma$ -equiv. classes



- Oval shapes: α -equivalence classes
- Each element $b \in B$ induces an equivalence relation α_b on C
- In this example, there are only two such equivalence relations (of form α_b). Not true in general!
- Can pp-define 3-ary relation

$$R = \{((b, c), (b_s, c_s), (b_t, c_t)) \mid (c_s, c_t) \in \alpha_b\}$$

$$R((v^b, v^c), (v_s^b, v_s^c), (v_t^b, v_t^c)) \equiv \exists y_s, y_t (\beta((v^b, v^c), y_s) \wedge \alpha(y_s, y_t) \wedge \gamma(y_s, (v_s^b, v_s^c)) \wedge \gamma(y_t, (v_t^b, v_t^c)))$$

- Have two equivalence relations (on C) in set $\{\alpha_b\}$
- Can pp-define 3-ary relation

$$R = \{((b, c), (b_s, c_s), (b_t, c_t)) \mid (c_s, c_t) \in \alpha_b\}$$

- Towards hardness: consider relations of the form

$$R_1 = \{((b_1^1, c_1^1), \dots, (b_m^1, c_m^1), (b_s, c_s), (b_t, c_t)) \mid (c_s, c_t) \in \alpha_{f^1(b_1^1, \dots, b_m^1)}\}$$

- For what functions f^1 can we pp-define R_1 ?
 R tells us that we can compute projections

- Towards hardness: consider relations of the form

$$R_1 = \{((b_1^1, c_1^1), \dots, (b_m^1, c_m^1), (b_s, c_s), (b_t, c_t)) \mid (c_s, c_t) \in \alpha_{f^1}(b_1^1, \dots, b_m^1)\}$$

- Now consider a second relation of similar form

$$R_2 = \{((b_1^2, c_1^2), \dots, (b_m^2, c_m^2), (b_s, c_s), (b_t, c_t)) \mid (c_s, c_t) \in \alpha_{f^2}(b_1^2, \dots, b_m^2)\}$$

- We can define the meet...

$$R_1(x_1, \dots, x_m, y, y') \wedge R_2(x_1, \dots, x_m, y, y')$$

- We can also define the join!

$$\exists z(R_1(x_1, \dots, x_m, y, z) \wedge R_2(x_1, \dots, x_m, z, y'))$$

- Up to associating together values $b, b' \in B$ such that $\alpha_b = \alpha_{b'}$, we can compute, starting from a set of variables, projections, meets, and joins on a two-element lattice
- We can efficiently simulate (in pp formulas) **monotone boolean formulas!**
(propositional formulas over $\{\wedge, \vee\}$, without quantification)
- To prove hardness of equivalence on the particular pentagon considered, we can appeal to the coNP-hardness of...
Problem: decide if two monotone boolean formulas are equivalent

- On a pentagon, in general, we may have $|\{\alpha_b\}| > 2\dots$
 \Rightarrow need to consider computational problems on formulas on larger lattices
- In general case, we are not handed a pentagon!
- We work to “reduce” to the case of pentagons
- We find a generic failure of congruence modularity in the $\mathcal{V}(\mathbb{A}_{\mathbf{B}})$ -free algebra on 4 generators
- We show that there is a sequence of efficiently pp-definable relations $\{D_i\}_{i \geq 1}$ such that if $(b_1, \dots, b_n) \in D_n$, then b_1, \dots, b_n are all contained in some pentagon
- This allows us to deal with “sets of pentagons”
We use a hardness result on equivalence of lattice words that “uniformizes” over sets of pentagons, to get hardness

Open questions



- Please get to work on the conjectures
- As far as we know, this is the first time computational hardness results have been derived based on the algebraic conditions studied. Are there any other applications of these ideas?
- Modulo the two conjectures, we classified the studied problems on *finite* structures. What about infinite structures?

