Testing expressibility is hard

Ross Willard

University of Waterloo, Canada

CP 2010
St. Andrews, Scotland
September 7, 2010

Outline

0. Apology
1. The Definition
2. The Problem
3. Our Solution
4. The Proof (hints)
5. Open Questions

Part 0 - Apology

I'm sorry

Part 1 - The Definition

Let Γ be a constraint language (always finite, on a finite domain D).

Let R be a relation (on D), say of arity $k > 0$.

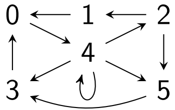
Definition

Γ can **express** R if there exist a CSP $\mathcal{P} = (X, D, \mathcal{C})$ over Γ (the **gadget**) and $x_1, \ldots, x_k \in X$ such that

$$\mathrm{proj}_{(x_1,\ldots,x_k)}(\mathrm{Sol}(\mathcal{P})) = R.$$

Example: $\Gamma = \{\rightarrow, U\}$, where

$D = \{0, 1, \ldots, 5\}$,



$U = \{0, 3\}$

Can Γ can express $R = \{3, 4, 5\}$?

Answer: Yes, via the gadget $\mathcal{P}$:

$$(a \rightarrow b) \ \& \ (b \rightarrow x) \ \& \ (x \rightarrow c) \ \& \ U(a) \ \& \ U(c).$$

Solutions to $\mathcal{P}$:

a	b	x	c
0	4	3	0
0	4	4	3
0	4	5	3
3	0	4	3

Part 2 - The Problem

Definition. $\textsc{Expr}(\Gamma)$ is the set of all relations expressible by Γ.

Expressibility Problem

Input:

- a constraint language Γ.
- a relation R.

Question:

- Is $R \in \textsc{Expr}(\Gamma)$?

Basic question: How hard is this problem to answer? To witness?

(Jeavons, Cohen, Gyssens; CP 1996, *Constraints* 1999):

Polymorphisms and 'Indicator problems'

1. If $R \notin \mathrm{Expr}(\Gamma)$, then this is witnessed by a **polymorphism** of Γ of a specified arity which does not preserve R.

2. If $R \in \mathrm{Expr}(\Gamma)$, then R is expressed by a canonical gadget called an **indicator problem**.

Hence we can (in principle) test whether $R \in \mathrm{Expr}(\Gamma)$ by either:

- Searching among all operations on D of a specified arity for one which witnesses $R \notin \mathrm{Expr}(\Gamma)$, **or**
- Building the indicator problem for (Γ, R), finding all of its solutions, and comparing the solution set to R.

How **practical** is this algorithm?

Answer: it's crap.

1. If $R \notin \text{EXPR}(\Gamma)$:
   general theory promises a witnessing polymorphism, but of arity $|R|$ (and hence table size $|D|^{|R|}$).

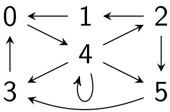
2. If $R \in \text{EXPR}(\Gamma)$:
   general theory offers the indicator problem for (Γ, R), but it has
   - $|D|^{|R|}$ variables
   - For each $S \in \Gamma$, $|S|^{|R|}$ constraints using S.

The size of the witness (polymorphism or gadget) guaranteed by theory is exponential in $|R|$. In particular, the JCG algorithm is co-*NEXPTIME*.

However: Practice does not always match theory.

Consider again $\Gamma = \{\rightarrow, U\}$:

$D = \{0, 1, \ldots, 5\}$,



$U = \{0, 3\}$

1. The relation $R = \{3, 4, 5\}$ **can** be expressed by Γ. Witnessing gadget:
   - Theory: $6^3 = 216$ variables and $10^3 + 2^3 = 1\,008$ constraints.
   - Practice: 4 variables and 5 constraints.
2. The relation $\nrightarrow \ := \ D^2 \setminus \rightarrow$ **cannot** be expressed by Γ. Witnessing polymorphism:
   - Theory: arity 26 (and thus table size $6^{26} \approx 1.7 \times 10^{20}$).
   - Practice: arity 1.

This is our problem:

- General theory seems to require witnesses of size $\exp(|R|)$.
- The Jeavons/Cohen/Gyssen algorithm is co-*NExpTime*.

But:

- Examples suggest that much smaller witnesses might suffice.
- Perhaps there is a better (polynomial-time?) algorithm for testing expressibility.

Positive evidence for better theory: boolean case ($D = \{0, 1\}$)

Theorem (Dalmau, 2000)

For each **fixed** *constraint language* Γ *over* $D = \{0, 1\}$, *testing membership in* $\textsc{Expr}(\Gamma)$ *is in P.*

Theorem (Creignou, Kolaitis & Zanuttini, 2008)

There is a uniform polynomial-time algorithm for testing expressibility over the boolean domain.

Consequences (boolean domain):

- Polynomial-sized gadgets if $R \in \textsc{Expr}(\Gamma)$;
- (Conjecture) $O(\log |R|)$-arity polymorphisms if $R \notin \textsc{Expr}(\Gamma)$.

What about the non-boolean case?

Conjecture (Vardi, AIM 2008): It's VERY VERY BAD. To wit,

- The general expressibility problem is co-*NExpTime*-complete, even on domain size 3.

(If true: no poly-sized witnesses in general; JCG algorithm can't be improved.)

Part 3 - Our solution

We confirm most of the AIM conjecture:

Theorem 1 (Exponential witnesses are unavoidable)

For infinitely many n there exist constraint languages Γ_0, Γ_1 and relation R, all over a 22-element domain, such that:

- $|R| = n$, and $\text{arity}(R) = O(\log n)$;
- $R \in \text{EXPR}(\Gamma_0)$, yet every witnessing gadget has $\geq 2^{n/3}$ variables;
- $R \notin \text{EXPR}(\Gamma_1)$, yet every witnessing polymorphism has arity $\geq n/3$.

Fine print: $|\Gamma_0| = |\Gamma_1| = O(n)$ and each $S \in \Gamma_0 \cup \Gamma_1$ has arity $O(\log n)$.

Theorem 2 (The JCG algorithm is essentially best possible)

There exists $d > 3$ such that testing expressibility on d-element domains is co-*NEXPTIME*-complete.

Part 4 - The Proof (hints)

1. Warning: it's complicated.

2. Following a suggestion of Vardi, we encode a class of **tiling problems** into the **complement** of the expressibility problem.

3. Did I mention that it's complicated?

Definition. A **tiling problem** is a special, succinctly presented CSP whose specification includes:

- A finite set Δ of **tile types**.
- A positive integer N, which determines an $N \times N$ grid.

$N = 5$

5	8	6	9	7
4	7	5	8	6
3	6	4	7	5
2	5	3	6	4
1	4	2	5	3

E.g.: $\Delta = \{1, 2, \ldots, 9\}$

Constraints:

H: $\boxed{s}\boxed{t} \quad \Rightarrow \quad t - s \in \{-2, 3\}$

V: $\begin{matrix}\boxed{t}\\ \boxed{s}\end{matrix} \quad \Rightarrow \quad s < t$

- **Constraints** on horizontally and vertically adjacent tile types.

Question: Can one cover the grid with tiles subject to the constraints?

- Optional input: An **initial condition** (prescribed tiles on first row).

Fix $\mathcal{D} = (\Delta, H, V)$.

- $\mathcal{D}$ is called a **domino system**.

"Exponential Tiling-by-$\mathcal{D}$ Problem," or $\textsc{ExpTile}(\mathcal{D})$

Input:

initial condition $\mathbf{w} \in \Delta^+$.

Question:

Can $\mathcal{D}$ tile the $2^m \times 2^m$ grid satisfying $\mathbf{w}$, where $m = |w|$?

Two facts about exponential tiling.

1. (Amusing exercise): there exists a domino system $\mathcal{D}_0$ such that:

   For all $m \geq m_0$ there exists $\mathbf{w}_m \in \Delta^m$ such that $\mathcal{D}_0$ can **almost tile** the $2^m \times 2^m$ grid satisfying initial condition $\mathbf{w}_m$ (i.e., cannot tile it, but can tile $N \times N$ for any $N < 2^m$).

   and

   For all $m \geq m_0$ there exists $\mathbf{x}_m \in \Delta^m$ such that $\mathcal{D}_0$ can tile the $2^m \times 2^m$ grid satisfying initial condition $\mathbf{x}_m$, but **cannot tile it with a repeated row**.

2. (Cf. Grädel *et al*) There exists a **universal** domino system $\mathcal{D}_1$ for *NEXPTIME*, i.e., such that $\textsc{ExpTile}(\mathcal{D}_1)$ is *NEXPTIME*-complete.

Key Construction of our paper

For every domino system $\mathcal{D} = (\Delta, H, V)$ there exists a finite set D and a log-space construction

$$\begin{array}{c} \mathbf{w} \in \Delta^+ \\ (|\mathbf{w}| = m) \end{array} \quad \mapsto \quad \left\{ \begin{array}{ccl} \Gamma & - & \text{constraint language with domain } D \\ R & - & \text{relation on } D \text{ of arity } k = O(\log m) \\ \mathbf{a} & - & \text{element in } D^k \setminus R \end{array} \right.$$

such that the **smallest** relation expressible from Γ and containing R . . .

- is either R or $R \cup \{\mathbf{a}\}$;
- is $R \cup \{\mathbf{a}\} \;\Leftrightarrow\; \mathcal{D}$ can tile $2^m \times 2^m$ with initial condition $\mathbf{w}$.

Hence $R \in \mathrm{EXPR}(\Gamma) \Leftrightarrow \mathcal{D}$ cannot tile $2^m \times 2^m$ with initial condition $\mathbf{w}$.

We also connect properties of tilings (‘almost,’ ‘repeated rows’) to possible sizes of witnesses to expressibility. (Theorems 1 & 2 follow.)

Part 5 - Questions we have not answered

1. Can non-existence of sub-exponential-sized witnesses be pushed down to smaller domains? To 3-element domains?

2. Can co-*NExpTime*-completeness be pushed down to small domains? To 3-element domains?

3. Does there exist a **fixed** constraint language Γ such that deciding membership in $\mathrm{E{\scriptstyle XPR}}(\Gamma)$:

   - Requires exponential-sized witnesses?
   - Is co-*NExpTime*-complete?

(Our construction $\mathbf{w} \mapsto (\Gamma, R, \mathbf{a})$ has Γ depending on $\mathbf{w}$.)

Thank you!