CSP, Algebras, Varieties

Andrei A. Bulatov
Simon Fraser University

CSP Reminder

An instance of CSP is defined to be a pair of relational structures A and B over the same vocabulary τ. Does there exist a homomorphism ϕ: A $\rightarrow$ B?

Example: Graph Homomorphism, H-Coloring

Given a sentence $\exists x_1, \ldots, x_n \quad R_1 \wedge \ldots \wedge R_k$ and a model for R_i decide whether or not the sentence is true

Example: SAT

CSP($\mathcal{B}$), CSP(B)

Polymorphisms Reminder

Definition

A relation (predicate) R is invariant with respect to an n-ary operation f (or f is a polymorphism of R) if, for any tuples $\vec{a}_1,\ldots,\vec{a}_n \in R$ the tuple obtained by applying f coordinate-wise is a member of R

Pol($\mathcal{A}$) denotes the set of all polymorphisms of relations of $\mathcal{A}$

Pol(Γ) denotes the set of all polymorphisms of relations from Γ

Inv(C) denotes the set of all relations invariant under operations from C

Outline

- From constraint languages to algebras
- From algebras to varieties
- Dichotomy conjecture, identities and meta-problem
- Datalog and variety
- Algebras and varieties in other constraint problems

Languages/polymorphisms vs. structures/algebras

Relational structure

$\mathcal{A} = (A, R_1, \ldots, R_m)$

$\Uparrow$

Constraint language

$\Gamma = \{R_1, \ldots, R_m\}$

Algebra

$\mathbf{A} = (A; f_1, \ldots, f_m)$

$\Uparrow$

Set of operations

$C = \{f_1, \ldots, f_m\}$

Inv(C) = Inv($\mathbf{A}$)

Pol(Γ) = Pol($\mathcal{A}$)

$Str(\mathbf{A})$ = (A; Inv($\mathbf{A}$))

$Alg(\mathcal{A})$ = (A; Pol($\mathcal{A}$))

CSP($\mathbf{A}$) = CSP($\mathcal{A}$)

Algebras - Examples

semilattice operation $\bullet$ semilattice $(A; \bullet)$

$$x \bullet x = x, \;\; x \bullet y = y \bullet x, \;\; (x \bullet y) \bullet z = x \bullet (y \bullet z)$$

affine operation $f(x,y,z)$ affine algebra $(A; f)$

$$f(x,y,z) = x - y + z$$

group operation $\bullet$ group $(A; \bullet, ^{-1}, 1)$

permutations $g_1, \ldots, g_k$ G-set $(A; g_1, \ldots, g_k)$

Expressive power and term operations

$\mathcal{A} = (A, R_1, \ldots, R_m)$

$\mathbf{A} = (A; f_1, \ldots, f_m)$

Expressive power

$\langle R_1, \ldots, R_m \rangle = \mathrm{InvPol}(R_1, \ldots, R_m)$

Term operations

$\langle f_1, \ldots, f_m \rangle = \mathrm{Pol}\,\mathrm{Inv}(f_1, \ldots, f_m)$

Primitive positive definability

$$R(x_1, \ldots, x_k) = \exists y_1, \ldots, y_n \; \Phi(x_1, \ldots, x_k, y_1, \ldots, y_n)$$

Substitutions

$$h(x_1, \ldots, x_k) = f(g_1(x_1, \ldots, x_k), \ldots, g_n(x_1, \ldots, x_k))$$

Clones of relations

Clones of operations

Good and Bad

Theorem (Jeavons)

Relational structure is good if *all* relations in its expressive power are good

Algebra is good if it has *some* good term operation

Relational structure is bad if *some* relation in its expressive power is bad

Algebra is bad if *all* its term operation are bad

Subalgebras

A set $B \subseteq A$ is a subalgebra of algebra $\mathbf{A} = (A; f_1, \ldots, f_m)$ if every operation of $\mathbf{A}$ preserves B

In other words $B \in \mathrm{Inv}(f_1, \ldots, f_m)$

It can be made an algebra $\mathbf{B} = (B; f_1|_B, \ldots, f_m|_B)$

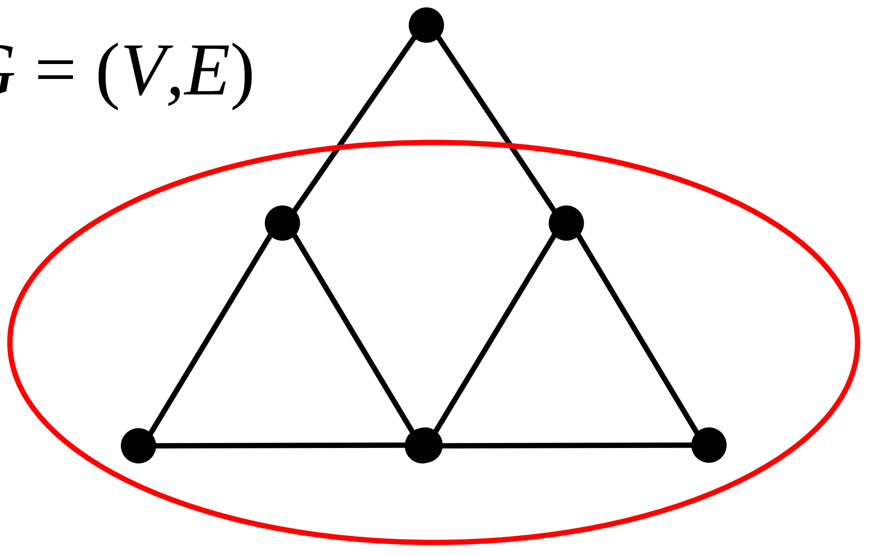
$(Z_4; x - y + z)$, $Z_4 = \{0,1,2,3\}$ $\{0,2\}$, $\{1,3\}$ are subalgebras

$\{0,1\}$ is not

$(Z_4; \max(x, y))$, any subset is a subalgebra

Subalgebras - Graphs

$G = (V,E)$



What subalgebras of $Alg(G)$ are?

$$G \rightarrow Alg(G) \rightarrow \mathrm{Inv}(Alg(G)) \rightarrow \mathrm{InvPol}(G)$$

$$B(x) = \exists y,z \ (E(x,y) \wedge E(y,z) \wedge E(z,x))$$

Subalgebras - Reduction

Theorem (B,Jeavons,Krokhin)

Let $\mathbf{B}$ be a subalgebra of $\mathbf{A}$. Then CSP(B) $\subseteq$ CSP(A)

Every relation $R \in$ Inv($\mathbf{B}$) belongs to Inv($\mathbf{A}$)

Take operation f of $\mathbf{A}$ and $(a_{11},\ldots,a_{1k}),\ldots,(a_{n1},\ldots,a_{nk}) \in R$

$$\begin{array}{c} f \\ \hline a_{11},\ldots,a_{1k} \\ \vdots \\ a_{n1},\ldots,a_{nk} \\ \hline b_1,\ldots,b_k \end{array} \quad \in R$$

Since $f|_B$ is an operation of $\mathbf{B}$

Homomorphisms

Algebras $\mathbf{A} = (A; f_1^A, \ldots, f_m^A)$ and $\mathbf{B} = (B; f_1^B, \ldots, f_m^B)$ are similar if f_i^A and f_i^B have the same arity

A homomorphism of $\mathbf{A}$ to $\mathbf{B}$ is a mapping φ: $A \to B$ such that

$$\varphi(f_i^A(x_n, \ldots, x_n)) = f_i^B(\varphi(x_1), \ldots, \varphi(x_n))$$

Homomorphisms - Examples

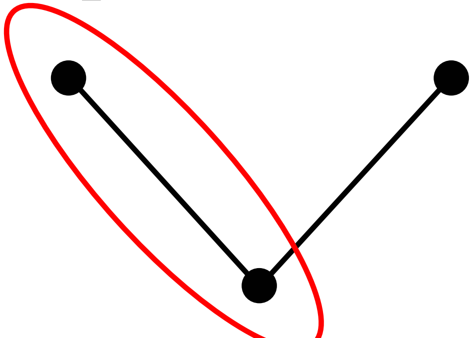
Affine algebras

$$(Z_4; x -_4 y +_4 z) \;\rightarrow\; (Z_2; x -_2 y +_2 z)$$

$$\varphi : \begin{array}{l} 0,2 \rightarrow 0 \\ 1,3 \rightarrow 1 \end{array}$$

Semilattices

$\mathbf{S}_1 = (\{a, b, c\}; \bullet^1)$



$\mathbf{S}_2 = (\{0,1\}; \bullet^2)$

Homomorphisms - Congruences

Let $\mathbf{B}$ is a homomorphic image of $\mathbf{A} = (A; f_1, \ldots, f_m)$ under homomorphism φ. Then the kernel of φ:

$(a,b) \in$ ker(φ) $\Leftrightarrow$ φ(a) = φ(b)

is a congruence of $\mathbf{A}$

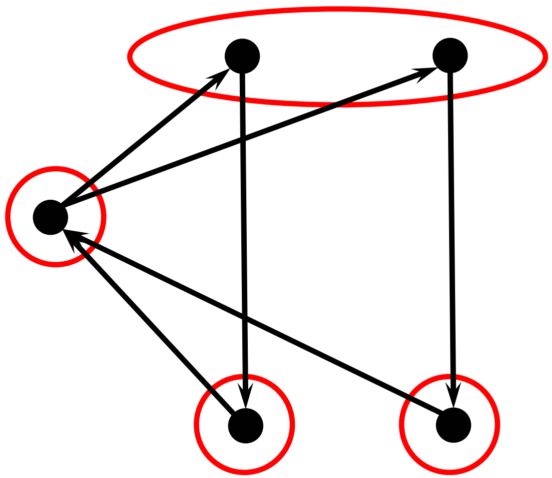
$$\begin{pmatrix} f^A(a_1, \ldots, a_n) \\ f^A(b_1, \ldots, b_n) \end{pmatrix} \in \ker(\varphi) \quad \Leftrightarrow$$

$$\Leftrightarrow \varphi(f^A(a_1, \ldots, a_n)) = \varphi(f^A(b_1, \ldots, b_n))$$

$$\Leftrightarrow f^A(\varphi(a_1), \ldots, \varphi(a_n)) = f^A(\varphi(b_1), \ldots, \varphi(b_n))$$

Congruences are equivalence relations from Inv($\mathbf{A}$)

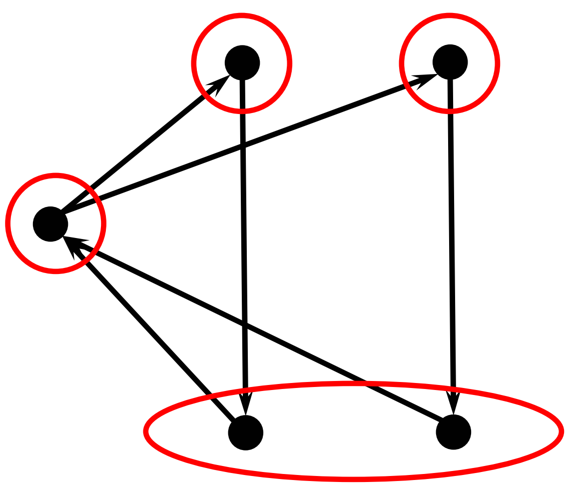
Homomorphisms - Graphs



$G = (V,E)$

$$\theta(x,y) = \exists z \ (E(z,x) \wedge E(z,y))$$

congruences of $Alg(G)$



$$\eta(x,y) = \exists z \ (E(x,z) \wedge E(y,z))$$

Homomorphisms - Reduction

Theorem

Let $\mathbf{B}$ be a homomorphic image of $\mathbf{A}$. Then for every finite $\Gamma \subseteq \text{Inv}(\mathbf{B})$ there is a finite $\Delta \subseteq \text{Inv}(\mathbf{A})$ such that CSP(Γ) is poly-time reducible to CSP(Δ)

Instance of CSP(Γ)

$$\bigwedge R(x_{i_1}, \ldots, x_{i_m})$$

$$\bigwedge \varphi^{-1}(R)(x_{i_1}, \ldots, x_{i_m})$$

Instance of CSP(Δ)

Direct Power

The nth direct power of an algebra $\mathbf{A} = (A; f_1, \ldots, f_m)$ is the algebra $\mathbf{A}^n = (A^n; f_1^n, \ldots, f_m^n)$ where the f_i^n act component-wise

$$f_i^n \begin{pmatrix} a_{11} & & a_{1k} \\ \vdots & ,\ldots, & \vdots \\ a_{n1} & & a_{nk} \end{pmatrix} = \begin{pmatrix} f_1(a_{11},\ldots,a_{1k}) \\ \vdots \\ f_1(a_{n1},\ldots,a_{nk}) \end{pmatrix}$$

Observation

An n-ary relation from Inv($\mathbf{A}$) is a subalgebra of $\mathbf{A}^n$

Direct Product - Reduction

Theorem

CSP($\mathbf{A}^n$) is poly-time reducible to CSP(A)

Transformations and Complexity

Theorem

Every subalgebra, every homomorphic image and every power of a tractable algebra are tractable

Corollary

If an algebra has an NP-complete subalgebra or homomorphic image then it is NP-complete itself

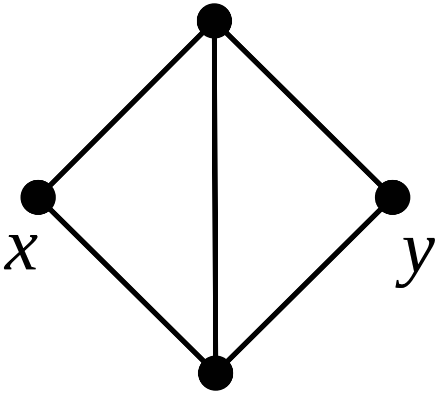
H-Coloring Dichotomy

Using G-sets we can prove NP-completeness of the
k-Coloring problem (or K_k-Coloring)

Take a non-bipartite graph H

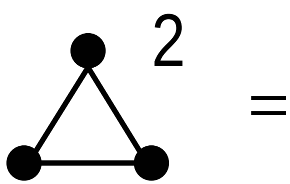
- Replace it with a subalgebra of all nodes in triangles
- Take homomorphic image modulo the transitive closure of the following

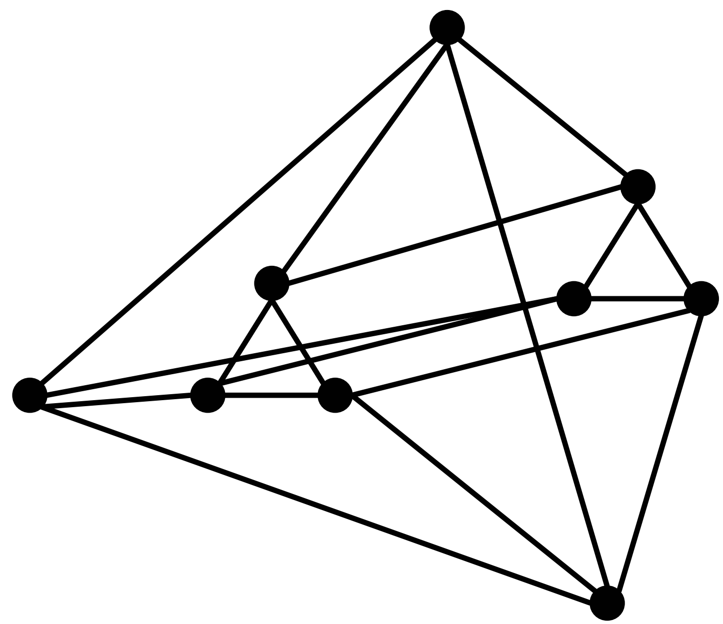
$\theta(x,y) =$



H-Coloring Dichotomy (Cntd)

- What we get has a subalgebra isomorphic to a power of a triangle





- It has a homomorphic image which is a triangle

This is a hom. image of an algebra, not a graph!!!

Varieties

Variety is a class of algebras closed under taking subalgebras, homomorphic images and direct products

Take an algebra $\mathbf{A}$ and built a class by including all possible direct powers (infinite as well), subalgebras, and homomorphic images.
We get the variety var($\mathbf{A}$) generated by $\mathbf{A}$

Theorem

If $\mathbf{A}$ is tractable then any finite algebra from var($\mathbf{A}$) is tractable

If var($\mathbf{A}$) contains an NP-complete algebra then $\mathbf{A}$ is NP-complete

Meta-Problem - Identities

Meta-Problem

Given a relational structure (algebra), decide if it is tractable

HSP Theorem A variety can be characterized by identities

Semilattice $x \bullet x = x,\ x \bullet y = y \bullet x,\ (x \bullet y) \bullet z = x \bullet (y \bullet z)$

Affine $\rightarrow$ Mal'tsev $f(x,y,y) = f(y,y,x) = x$

Near-unanimity

$$f(y,x,\ldots,x) = f(x,y,\ldots,x) = \ldots = f(x,\ldots,x,y) = x$$

Constant $f(x_1,\ldots,x_k) = f(y_1,\ldots,y_k)$

Dichotomy Conjecture - Identities

Dichotomy Conjecture

A finite algebra $\mathbf{A}$ is tractable if and only if var($\mathbf{A}$) has a Taylor term: $t(z_1,\ldots,z_n)$

$$t(x_{i1},\ldots,x_{in}) = t(y_{i1},\ldots,y_{in}),\ x_{ii} \neq y_{ii} \text{ for } i = 1\ldots n$$

Otherwise it is NP-complete

Idempotent algebras

An algebra is called surjective if every its term operation is surjective

If a relational structure $\mathbb{A}$ is a core then $Alg(\mathbb{A})$ is surjective

Theorem

It suffices to study surjective algebras

Theorem

It suffices to study idempotent algebras: $f(x,\ldots,x) = x$

If $\mathbf{A}$ is idempotent then a G-set belongs to var($\mathbf{A}$) iff it is a divisor of $\mathbf{A}$, that is a hom. image of a subalgebra

Dichotomy Conjecture

Dichotomy Conjecture

A finite idempotent algebra $\mathbf{A}$ is tractable if and only if var($\mathbf{A}$) does not contain a finite G-set. Otherwise it is NP-complete

This conjecture is true if

- $\mathbf{A}$ is a 2-element algebra (Schaefer)
- $\mathbf{A}$ is a 3-element algebra (B.)
- $\mathbf{A}$ is a conservative algebra, (B.)
  i.e. every subset of its universe is a subalgebra

Complexity of Meta-Problem

Theorem (B.,Jeavons)

- The problem, given a finite relational structure $\mathcal{A}$, decide if $Alg(\mathcal{A})$ generates a variety with a G-set, is NP-complete

- For any k, the problem, given a finite relational structure $\mathcal{A}$ of size at most k, decide if $Alg(\mathcal{A})$ generates a variety with a G-set, is poly-time

- The problem, given a finite algebra $\mathbf{A}$, decide if it generates a variety with a G-set, is poly-time

Definability in Datalog

Theorem (Feder,Vardi)

If a relational structure $\mathcal{A}$ is such that Pol($\mathcal{A}$) contains a near-unanimity operation then $\overline{\text{CSP}(\mathcal{A})}$ is definable in Datalog

If a relational structure $\mathcal{A}$ is such that Pol($\mathcal{A}$) contains a semilattice operation then $\overline{\text{CSP}(\mathcal{A})}$ is definable in Datalog

Datalog and Algebras

Theorem (Larose, Zadori)

Let $\mathcal{A}$, $\mathcal{A}'$ be finite relational structures with the same universe and such that Pol($\mathcal{A}$) $\subseteq$ Pol($\mathcal{A}'$). Then if $\overline{\text{CSP}(\mathcal{A})}$ is definable in Datalog then so is $\overline{\text{CSP}(\mathcal{A}')}$

It makes sense to talk about algebras with $\overline{\text{CSP}(\mathbf{A})}$ definable in Datalog

Theorem

If $\overline{\text{CSP}(\mathbf{A})}$ is definable in Datalog then so is $\overline{\text{CSP}(\mathbf{B})}$ for any algebra $\mathbf{B}$ from var($\mathbf{A}$)

Linear Equations

Linear Equation:

Is a system of linear equations over a finite field consistent?

$$\begin{cases} 2x+3y=5 \\ x+y-2z=-1 \end{cases}$$

CSP form: $\exists x,y,z \quad (2x+3y=5)\wedge(x+y-2z=-1)$

Linear Equation is equivalent to CSP($\mathbf{A}$) where

$$\mathbf{A} = (A;\ x - y + z)$$

Fact Linear Equation is not definable in Datalog

Conjecture and Meta-Problem for Datalog

Theorem

If var($\mathbf{A}$) contains a G-set or an affine algebra then $\overline{\mathrm{CSP}(\mathbf{A})}$ is not definable in Datalog

Conjecture

$\overline{\mathrm{CSP}(\mathbf{A})}$ is definable in Datalog iff var($\mathbf{A}$) does not contain a G-set or an affine algebra

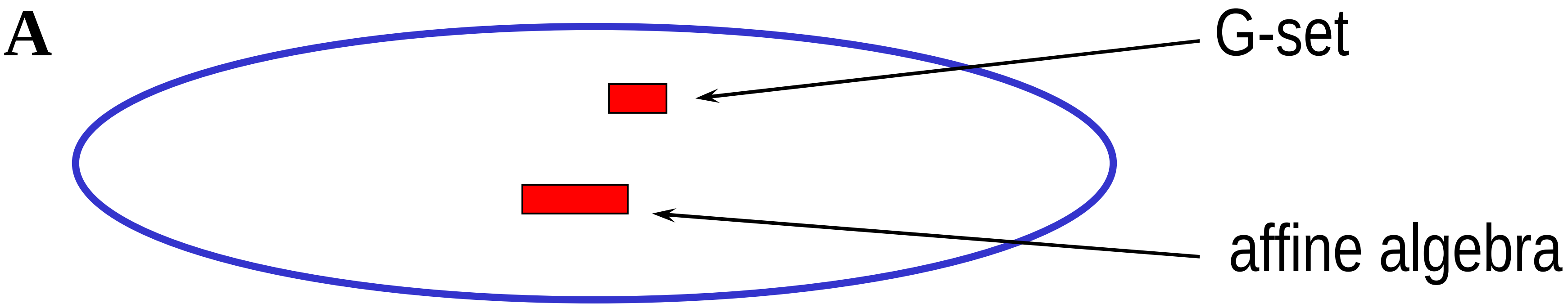
The Conjecture is true for 2-, 3-element and conservative algebras

Complexity of Meta-Problem for Datalog

Theorem

- The problem, given a finite relational structure $\mathcal{A}$, decide if $Alg(\mathcal{A})$ generates a variety with a G-set, or an affine algebra is NP-complete
- For any k, the problem, given a finite relational structure $\mathcal{A}$ of size at most k, decide if $Alg(\mathcal{A})$ generates a variety with a G-set or an affine algebra, is poly-time
- The problem, given a finite algebra $\mathbf{A}$, decide if it generates a variety with a G-set or an affine algebra, is poly-time

Localization - Types



Possible local structure:

1. G-set
2. linear space
3. 2-element Boolean algebra
4. 2-element lattice
5. 2-element semilattice

Omitting Types

Conjectures

$\mathbf{A}$ is tractable iff var($\mathbf{A}$) omits type 1

$\overline{\text{CSP}(\mathbf{A})}$ is definable in Datalog iff var($\mathbf{A}$) omits types 1 and 2

Other Problems: Counting

In a #CSP, given structures $\mathcal{A}$ and $\mathcal{B}$, we are asked how many homomorphisms from $\mathcal{A}$ to $\mathcal{B}$ are there.

Can be parametrized by relational structures in the usual way

Can be parametrized by algebras AND varieties

For #CSP(**A**) a classification is also known (B.,Dalmau,Grohe):

#CSP(**A**) is solvable in poly-time iff var(**A**) has a Mal'tsev term $f(x,y,y) = f(y,y,x) = x$ and every finite algebra from var(**A**) satisfies a certain condition on congruences

Other Problems: QCSP

In a QCSP we need decide if a positive conjunctive sentence is true in a given model

Can be parametrized by relational structures in the usual way

Complexity classification is almost known for 3-element structures (Chen),
and known for conservative structures (Feder/Chen)

QCSPs can be parametrized by algebras, but not varieties!

No feasible di/trichotomy conjecture known