Constraint Satisfaction Problems and Algebra

A mountaineer's-eye view

Benoît Larose

Department of Mathematics and Statistics
Concordia University, Montréal

Department of Mathematics
Champlain Regional College
St-Lambert QC, Canada

Workshop on Finite and Algorithmic Model Theory
May 14-18, 2012
Les Houches, France

Overview

In the last 15 years, universal algebraic techniques have led to deep insights into the nature of the computational and descriptive complexity of Constraint Satisfaction Problems. We present an overview of these techniques, as well as results and open questions in this area.

Part I: Base Camp

3 faces of CSPs

- **Variable-value**
  Given finite sets V (variables), D (values), and a set of constraints $\{(\overline{s}_1, R_1), \ldots, (\overline{s}_q, R_q)\}$ over V, is there a function $\varphi : V \to D$ such that $\varphi(\overline{s}_i) \in R_i$ for all i?
- **Conjunctive Query**
  Given a finite relational structure $\mathbb{B} = (D; \ R_1^{\mathbb{B}}, \ldots, R_k^{\mathbb{B}})$, and a primitive positive (pp) sentence

  $$\phi = \exists v_1 \ldots \exists v_n \ (R_1(v_{i_1}, \ldots, v_{i_k}) \wedge \ldots)$$

  does $\mathbb{B} \models \phi$ hold ?
- **Homomorphism**
  Given two similar relational structures, $\mathbb{A} = (V; \ R_1^{\mathbb{A}}, \ldots, R_k^{\mathbb{A}})$ and $\mathbb{B} = (D; \ R_1^{\mathbb{B}}, \ldots, R_k^{\mathbb{B}})$, is there a homomorphism $h : \mathbb{A} \to \mathbb{B}$?

CSPs: Homomorphism Form

- Given two similar relational structures, $\mathbb{A} = (V;\ R_1^{\mathbb{A}}, \ldots, R_k^{\mathbb{A}})$ and $\mathbb{B} = (D;\ R_1^{\mathbb{B}}, \ldots, R_k^{\mathbb{B}})$, is there a *homomorphism* $h : \mathbb{A} \to \mathbb{B}$?

  $\forall i\ [(a_1, \ldots, a_n) \in R_i^{\mathbb{A}} \Longrightarrow (h(a_1), \ldots, h(a_n)) \in R_i^{\mathbb{B}}]$

- If $\mathcal{A}$ and $\mathcal{B}$ are classes of similar structures:

$$\mathrm{CSP}(\mathcal{A}, \mathcal{B}) = \{(\mathbb{A}, \mathbb{B}) \mid \mathbb{A} \to \mathbb{B}, \mathbb{A} \in \mathcal{A}, \mathbb{B} \in \mathcal{B}\}$$

Fixed-target CSPs

- Universal algebraic techniques have been successful in the case of fixed-target CSPs, i.e.

$$\mathrm{CSP}(\mathcal{A}\mathit{ll}, \mathbb{B})$$

- We denote this by $\mathrm{CSP}(\mathbb{B})$ or $\mathrm{Hom}(\mathbb{B})$, so that

$$\begin{aligned} \mathrm{CSP}(\mathbb{B}) &= \{\mathbb{A} \mid \mathbb{A} \to \mathbb{B}\} \\ \neg\mathrm{CSP}(\mathbb{B}) &= \{\mathbb{A} \mid \mathbb{A} \not\to \mathbb{B}\} \end{aligned}$$

- The case $\mathrm{CSP}(\mathcal{B}, \mathcal{A}\mathit{ll})$ has been essentially settled by Grohe, 2003 (+ Dalmau, Kolaitis, Vardi, 2002)

Core structures

If $\mathbb{B}$ and $\mathbb{B}_0$ are homomorphically equivalent, i.e.
if $\mathbb{B}_0 \to \mathbb{B}$ and $\mathbb{B} \to \mathbb{B}_0$, then

$$CSP(\mathbb{B}_0) = CSP(\mathbb{B}).$$

Hence we may always assume $\mathbb{B}$ is a **core**,
i.e.
every homomorphism from $\mathbb{B}$ to $\mathbb{B}$ is onto,
i.e.
of all structures equivalent to $\mathbb{B}$, $\mathbb{B}$ has smallest universe.

Examples

- Many standard problems complete for important complexity classes such as **L**, **NL**, $\oplus$**L**, **P**, **NP** are of the form $\mathrm{CSP}(\mathbb{B})$:
- 2-SAT, 3-SAT, HORN 3-SAT, k-COL, systems of linear equations over $\mathbb{F}_p$, (un)directed unreachability, etc.

2-SAT: $\mathbb{B} = \langle \{0,1\}; \theta_{(x \vee y)}, \theta_{(x \vee \overline{y})}, \theta_{(\overline{x} \vee y)}, \theta_{(\overline{x} \vee \overline{y})} \rangle$ where

$$\theta_{(x \vee y)} = \{(0,1),(1,0),(1,1)\},$$
$$\theta_{(x \vee \overline{y})} = \{(0,0),(1,0),(1,1)\},$$

etc.

HORN-3-SAT: $\mathbb{B} = \langle \{0,1\}; \{0\}, \{1\}, \rho \rangle$ where

ρ is $\{(x,y,z) : (y \wedge z) \rightarrow x\}$;

Lin. Eq. mod p: $\mathbb{B} = \langle \{0,\ldots,p-1\}; \rho, \{b\} \rangle$ where

$b \neq 0$ and ρ is $\{(x,y,z) : x+y=z\}$;

Dir-Unreachability: $\mathbb{B} = \langle \{0,1\}; \leq, \{0\}, \{1\} \rangle$ where

$\leq$ is $\{(0,0),(0,1),(1,1)\}$;

2-colouring: $\mathbb{B} = \langle \{0,1\}; E \rangle$ where $E = \{(0,1),(1,0)\}$.

Classification Problems

Two main classification problems for $\mathrm{CSP}(\mathbb{B})$:

1. Classify $\mathrm{CSP}(\mathbb{B})$ w.r.t. *computational complexity*, i.e., w.r.t. membership in a given complexity class (e.g. **P**, **NL**, **L**), modulo assumptions like **P** $\neq$ **NP**;
2. Classify $\mathrm{CSP}(\mathbb{B})$ w.r.t. *descriptive complexity*, i.e., w.r.t. definability of $\mathrm{CSP}(\mathbb{B})$ in a given logic (FO, Datalog and its fragments - linear, symmetric)

The Dichotomy Conjecture

In particular:

Conjecture (Feder, Vardi, 1993)

Every $\mathrm{CSP}(\mathbb{B})$ *is either in* **P** *or* **NP***-complete.*

- Motivation: Feder and Vardi looking for a largest (synctactic) subclass of **NP** where no problems of intermediate complexity can be found (cf Ladner);
- Every problem in the suggested subclass is poly-time equivalent to some $\mathrm{CSP}(\mathbb{B})$ (FV 1998 + Kun 2006)

Dichotomy Conjecture, cont'd

"The importance of the Dichotomy Conjecture may not be so much what it literally says, namely that all CSPs are in **P** *or* **NP***-complete and that there is nothing in between, but the likely event that a satisfactory proof would provide deep algorithmic understanding of all tractable CSPs. (...)"*

A. Atserias

Two Natural Families of Tractable CSPs

1. *Few Subpowers*
2. *Datalog*

Outline of the remainder of the talk:

- II
  - A Just enough algebra to get us started;
  - B Connecting Algebras and CSPs;
  - C Algebras with Few Subpowers.
- III
  - A Datalog and friends;
  - B A bit more Algebra (TCT and Reductions);
  - C CSP's definable in Datalog and fragments.

Part II: Promenade

II A. Basic Universal Algebra

Let U be a non-empty set.

- $f : U^k \to U$: a k-ary operation on U.
- An **algebra**: $\mathbf{A} = \langle U; F \rangle$: F a set of operations on *universe* U.
- A **term**: any formal expression built from (names for) fundamental operations and variables;
- terms in k-variables define the k-ary **term operations** of $\mathbf{A}$.
- An operation f is **idempotent** if $f(x, \ldots, x) = x \ \forall x \in U$;
- $\mathbf{A}$ is **idempotent** if $\forall f \in F$ ($\iff$ $\forall$ term operation) f idempotent.
- In the sequel, all algebras will be assumed idempotent.

A class of algebras is

- **equational** if it can be axiomatised by **identities**, i.e. (universally quantified) equations between terms;
- a **variety** if it is closed under forming (H) homomorphic images, (S) subalgebras and (P) products.

Theorem

1. *(G. Birkhoff) Varieties = Equational classes.*
2. *(Tarski) The smallest variety* $var(\mathcal{K})$ *containing* $\mathcal{K}$ *is* $var(\mathcal{K}) = HSP(\mathcal{K})$.

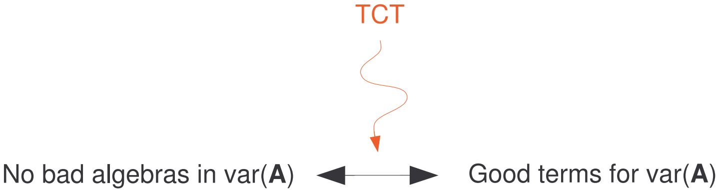
Example

The class of all algebras $\mathbf{S} = \langle S; \wedge^{\mathbf{S}} \rangle$ satisfying the identities

$$\begin{aligned} x \wedge (y \wedge z) &\approx (x \wedge y) \wedge z, \\ x \wedge y &\approx y \wedge x, \\ x \wedge x &\approx x, \end{aligned}$$

is the (equational) class of semilattices. It coincides with $var(\mathbf{T})$, where $\mathbf{T} = \langle \{0, 1\}; \wedge \rangle$ denotes the 2-element (meet) semilattice.
In other words, every semilattice is a homomorphic image of a subalgebra of a power of $\mathbf{T}$.

- Tame Congruence Theory:
- developed mid-80's by Hobby McKenzie and collaborators:
- connects local behaviour of finite algebras in (locally finite) varieties and existence of terms satisfying certain "good" identities;



II B. Connecting Relational Structures and Algebras

Let U be a finite set.

- Let f be a k-ary operation on U;
- Let $\theta \subseteq U^n$ be an n-ary relation on U.
- The operation f **preserves** the relation θ, or θ is **invariant** under f, if the following holds:

$$\underbrace{\begin{bmatrix} u_{1,1} & \cdots & u_{1,k} \\ \vdots & \cdots & \vdots \\ u_{n,1} & \cdots & u_{n,k} \end{bmatrix}}_{\text{columns in } \theta} \stackrel{f}{\mapsto} \underbrace{\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}}_{\theta}$$

Applying f to rows of the matrix with columns in θ yields a tuple of θ,

i.e. $f : \langle U; \theta \rangle^k \to \langle U; \theta \rangle$ is a homomorphism.

The algebra $\mathbf{A}(\mathbb{B})$

- Let $\mathbb{B} = \langle U; R_1, \ldots \rangle$ be a (core) relational structure.
- **polymorphisms** of $\mathbb{B}$: operations f on U that preserve every basic relation R_i of $\mathbb{B}$.

Definition

The (idempotent) algebra associated to $\mathbb{B}$ is

$$\mathbf{A}(\mathbb{B}) = \langle U; F \rangle$$

where F is the set of all *idempotent* polymorphisms of $\mathbb{B}$.

Subpowers

Subpower of $\mathbf{A}(\mathbb{B})$: (subuniverse) subalgebra of a finite power of $\mathbf{A}(\mathbb{B})$.

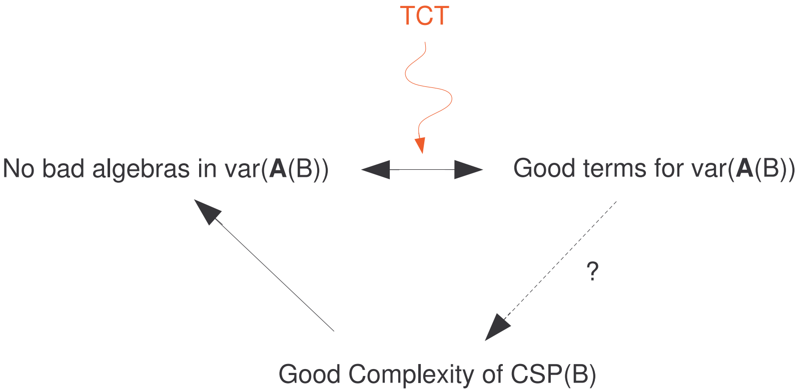
Theorem (Bodnarchuk et al, Geiger 60's)

Let $\mathbb{B} = \langle U; R_1, \ldots \rangle$. *Tfae:*

1. θ *is a subpower of* $\mathbf{A}(\mathbb{B})$*;*
2. θ *is invariant under every basic operation of* $\mathbf{A}(\mathbb{B})$*;*
3. θ *is pp-definable from the* R_i*.*

The Paradigm

The equational properties of $var(\mathbf{A}(\mathbb{B}))$ *control the descriptive and algorithmic complexity of the decision problem* $CSP(\mathbb{B})$.



On nice terms

Some CSP-friendly terms (all are idempotent):

- **semilattice**, e.g. min, max;
- **Maltsev**: $m(y,x,x) = m(x,x,y) = y$, e.g. $xy^{-1}z$ for a group;
- k-ary **weak NU**: $f(y,x,\ldots,x) = ... = f(x,\ldots,x,y)$ $(k \geq 2)$;
- k-ary **NU**: $f(y,x,\ldots,x) = \cdots = f(x,\ldots,x,y) = x$ $(k \geq 3)$;
- **Cyclic**: $f(x_1,\ldots,x_k) = \cdots = f(x_k,x_1,\ldots,x_{k-1})$ $(k \geq 2)$;
- **TSI term**: $f(x_1,\ldots) = f(y_1,\ldots)$ if $\{x_1,\ldots\} = \{y_1,\ldots\}$ $(k \geq 2)$;
- **Cube term**: for each position i, $\exists$ identity

$$f(\ldots,\underbrace{y}_{i},\ldots) = x \quad (k \geq 3);$$

II C: A First Application: Algebras with Few Subpowers

"This raises an important question. For subgroup problems, the set of solutions has a polynomial number of generators. This means that there exists a polynomial number of solutions such that if we take the closure under the mapping (...) then we obtain all solutions. (...)"

T. Feder, M. Vardi

The **subuniverse of A generated by** $X \subseteq U$: smallest subuniverse of **A** containing X. Subuniverse *k-generated*: has generating set of size $\leq k$.

Definition

Let **A** be a finite algebra.

- Let $g_{\mathbf{A}}(n) =$ least k such that every subuniverse of $\mathbf{A}^n$ is k-generated.

Definition

The algebra **A** has **polynomially generated subpowers (PGS)** if $\exists k \; g_{\mathbf{A}}(n) \in \mathcal{O}(n^k)$.

- Groups, rings, vector spaces: have PGS;
- Maltsev term $\implies$ PGS;
- NU term $\implies$ PGS;
- Semilattice term: NO

Theorem (BIMMVW 2010)

The following are equivalent:

1. **A** *has few subpowers;*
2. **A** *has polynomially generated subpowers;*
3. **A** *has a* cube term.

- $\mathbf{A}(\mathbb{B})$ has few subpowers $\iff$ at most exponentially many n-ary relations pp-definable from the basic relations of $\mathbb{B}$;
- Origin of PGS: Feder-Vardi 1998;
- PGS formalised in Dalmau 2002; conjectured CSP is tractable;
- 2006 Dalmau's algorithm: works for CSPs with Maltsev or NU term.

Theorem (IMMVW 2010)

If $\mathbf{A}(\mathbb{B})$ *has few subpowers, then* $\mathrm{CSP}(\mathbb{B})$ *is tractable.*

- PGS $=$ CSPs solvable by Dalmau's algorithm;
- akin to Gaussian elimination;
- Idea:
  - sets of solutions are subpowers;
  - *compact representations* generate the sets of solutions via good terms;
  - starting with no constraints and adding one at a time, recompute compact representations of the solution set.

Part III A: Bivouac

III A: Datalog

- A *Datalog Program* consists of *rules*, and takes as input a relational structure:

$$\theta_1(x,y)\ :-\ \theta_2(w,u,x),\theta_3(x),R_1(x,y,z),R_2(x,w)$$

- relations R_1 and R_2: basic relations from the input structures (EDBs);
- relations θ_i: auxiliary relations (IDBs);
- the rule stipulates that if the condition on the righthand side (the *body* of the rule) holds, then the condition of the left (the *head*) should also hold.

- Let $\mathbb{B} = \langle\{0,1\}; E = \{(0,1),(1,0)\}\rangle$: $\mathrm{CSP}(\mathbb{B})$ is 2-COL;
- A Datalog program that accepts precisely non-bipartite graphs: uses a single binary (IDB) *OddPath*.

$$\begin{aligned} OddPath(x,y) &:- \quad E(x,y) \\ OddPath(x,y) &:- \quad OddPath(x,z), E(z,u), E(u,y) \\ \gamma &:- \quad OddPath(x,x) \end{aligned}$$

The 0-ary relation γ: *goal predicate* of the program: "lights up" precisely if the input structure admits NO homomorphism to the target structure $\mathbb{B}$.

Fragments of Datalog

A Datalog program is **linear** if each rule contains at most one occurrence of an IDB in the body, i.e. if each rule looks like this

$$\theta_1(x, y) \; :- \; \theta_2(w, u, x), R_1(x, y, z), R_2(x, w)$$

where the θ_i's are the only IDBs in it.

A linear Datalog program is **symmetric** if it is invariant under symmetry of rules, i.e. if the program contains the above rule, then it must also contain its **symmetric**:

$$\theta_2(w, u, x) \; :- \; \theta_1(x, y), R_1(x, y, z), R_2(x, w).$$

A Datalog program is **non-recursive** if no IDB appears in the body of any rule.

Definability in Datalog (and fragments)

$\neg\mathrm{CSP}(\mathbb{B})$ is **definable in (linear, symmetric, non-recursive) Datalog** if there exists a (linear, symmetric, non-recursive) Datalog program that accepts precisely $\neg\mathrm{CSP}(\mathbb{B})$.

- definability in Datalog $\iff$ $\exists posFO + LFP$;
- definability in non-recursive Datalog $\iff$ $\exists posFO$;

Notice that $\neg\mathrm{CSP}(\mathbb{B})$ is homomorphism-closed:

- (Atserias, 2008) $\neg\mathrm{CSP}(\mathbb{B})$ FO-definable $\iff$ $\exists posFO$;
- (Rossman 2008) for hom-closed classes, FO-definable $\iff$ $\exists posFO$;
- (Dawar, Kreutzer 2008) for hom-closed classes, $LFP \neq Datalog$.

Some facts about the complexity of CSPs in Datalog:

- $\neg\mathrm{CSP}(\mathbb{B})$ definable in Datalog $\Rightarrow \mathrm{CSP}(\mathbb{B}) \in$ **P**;
- $\neg\mathrm{CSP}(\mathbb{B})$ definable in linear Datalog $\Rightarrow \mathrm{CSP}(\mathbb{B}) \in$ **NL**;
- $\neg\mathrm{CSP}(\mathbb{B})$ definable in symmetric Datalog $\Rightarrow \mathrm{CSP}(\mathbb{B}) \in$ **L**.

The converse of the last two statements holds for all CSPs known to belong to **NL** and **L**.

Datalog and CSPs: Main Questions

- Feder Vardi: tractability of many CSPs explained by Datalog,
- but not all: Linear Equations.
- FV formulate a conjecture: (*Not* Datalog $\iff$ *Ability to Count*);

Some Questions about Datalog and CSPs

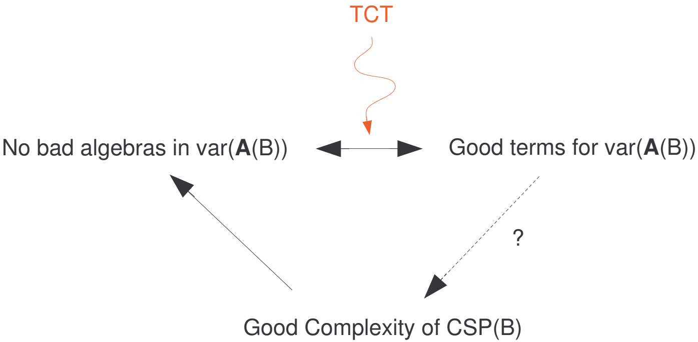
1. *For which* $\mathbb{B}$ *is* $\neg\mathrm{CSP}(\mathbb{B})$ *definable in (linear, symmetric, non-recursive) Datalog ?*
2. *For CSPs, do we have LFP $\neq$ Datalog ?*

Part III B: Ascent

III B: Algebras and Complexity

The Paradigm

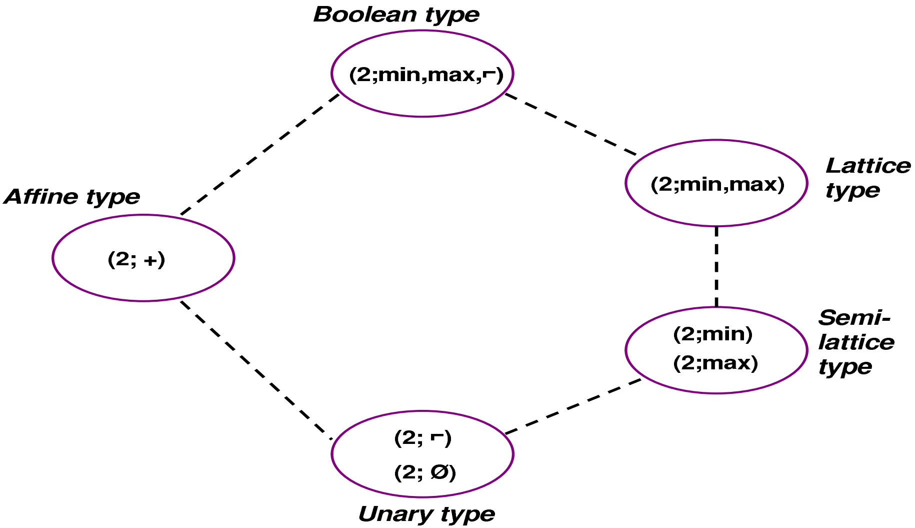
the equational properties of var$(\mathbf{A}(\mathbb{B}))$ *control the descriptive and algorithmic complexity of the decision problem CSP*$(\mathbb{B})$.



The "bad" algebras: typesets

- to each (finite) algebra $\mathbf{A}$ is associated a set of *types*;
- certain "prime" algebras have a *unique* type;
- the possible types are:
  - the *unary type*, or type 1;
  - the *affine* type, or type 2;
  - the *Boolean* type, or type 3;
  - the *lattice* type, or type 4;
  - the *semilattice* type, or type 5.
- the *typeset* of the variety $var(\mathbf{A})$ is the union of all typesets of all finite algebras in it;
- $var(\mathbf{A})$ **admits** or **omits** type i if i is or isn't in its typeset;

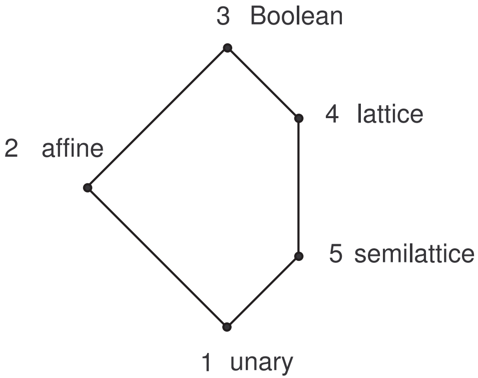
Boolean Functions and the Ordering of Types



Ordering of Types

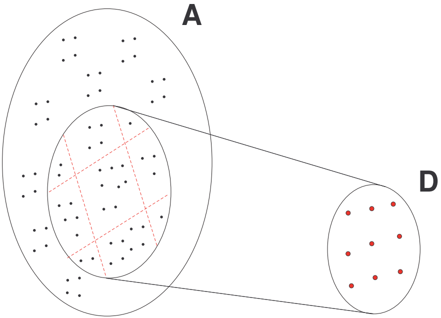
We refer later to the following ordering of types:

$$1 < 2 < 3 > 4 > 5 > 1$$



Definition

We say that the algebra **D** is a **divisor** of the algebra **A** if $\mathbf{D} \in HS(\mathbf{A})$, i.e. it is a homomorphic image of a subalgebra of **A**.



Divisors and Complexity

Lemma (Valeriote, 2007)

Let $\mathbf{A}$ *be an idempotent algebra, and suppose type* i *is in the typeset of* $var(\mathbf{A})$. *Then* $\mathbf{A}$ *has a ("prime") divisor of type* $\leq i$.

Lemma (BL, Tesson, 2007)

Let $\mathbb{B}$ *be a core. Let* $\mathbf{D}$ *be a divisor of* $\mathbf{A}(\mathbb{B})$, *and let* $\mathbb{B}'$ *be a structure whose basic relations are * invariant under the operations of* $\mathbf{D}$. *Then*

1. *there is an FO reduction ** of* $\mathrm{CSP}(\mathbb{B}')$ *to* $\mathrm{CSP}(\mathbb{B})$;
2. *if* $\neg\mathrm{CSP}(\mathbb{B})$ *is expressible in (Linear, Symmetric) Datalog then so is* $\neg\mathrm{CSP}(\mathbb{B}')$.

* ... irredundant and ...

** without ordering

- Classification of “prime” algebras (by type): Szendrei, 1992
- if **D** is “prime” of type i, its term operations preserve the basic relations of the structure $\mathbb{B}'$ where $\mathrm{CSP}(\mathbb{B}')$ is
  - 3-SAT if $i = 1$ (unary);
  - Lin. Eq. mod p if $i = 2$ (affine);
  - Dir. Unreachability if $i = 4$ (lattice);
  - HORN 3-SAT if $i = 5$ (semilattice).

Hardness results

Corollary (1)

Let $\mathbb{B}$ *be a core, and let* $\mathbf{A} = \mathbf{A}(\mathbb{B})$.

1. *(BJK, 2000) If var*$(\mathbf{A})$ *admits the unary type, then* $\mathrm{CSP}(\mathbb{B})$ *is* **NP***-complete;*
2. *if var*$(\mathbf{A})$ *admits the affine type, then* $\mathrm{CSP}(\mathbb{B})$ *is mod*$_p$**L***-hard* ($\exists p$)*;*
   Otherwise:
4. *if var*$(\mathbf{A})$ *admits the lattice type, then* $\mathrm{CSP}(\mathbb{B})$ *is* **NL***-hard;*
5. *if var*$(\mathbf{A})$ *admits the semilattice type, then* $\mathrm{CSP}(\mathbb{B})$ *is* **P***-hard.*

Non-expressibility results

Corollary (2)

Let $\mathbb{B}$ *be a core, and let* $\mathbf{A} = \mathbf{A}(\mathbb{B})$.

- *(BL, Zádori, 2006) If var*$(\mathbf{A})$ *admits the unary or affine type, then* $\neg\mathrm{CSP}(\mathbb{B})$ *is not expressible in Datalog;*
- *if var*$(\mathbf{A})$ *admits the semilattice type, then* $\neg\mathrm{CSP}(\mathbb{B})$ *is not expressible in Linear Datalog;*
- *if var*$(\mathbf{A})$ *admits the lattice type, then* $\neg\mathrm{CSP}(\mathbb{B})$ *is not expressible in Symmetric Datalog.*

Theorem

Let $\mathbb{B}$ *be a core. If* $\neg\mathrm{CSP}(\mathbb{B})$ *is not in* $\exists$*posFO then* $\mathrm{CSP}(\mathbb{B})$ *is* **L**-*hard.*

Algebraic Dichotomy Conjectures

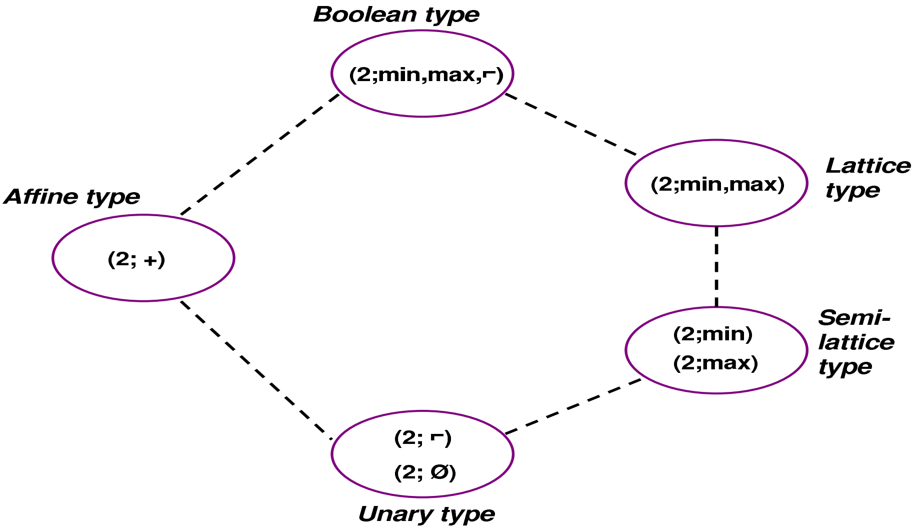
Conjecture

Let $\mathbb{B}$ be a core, and let $\mathbf{A} = \mathbf{A}(\mathbb{B})$.

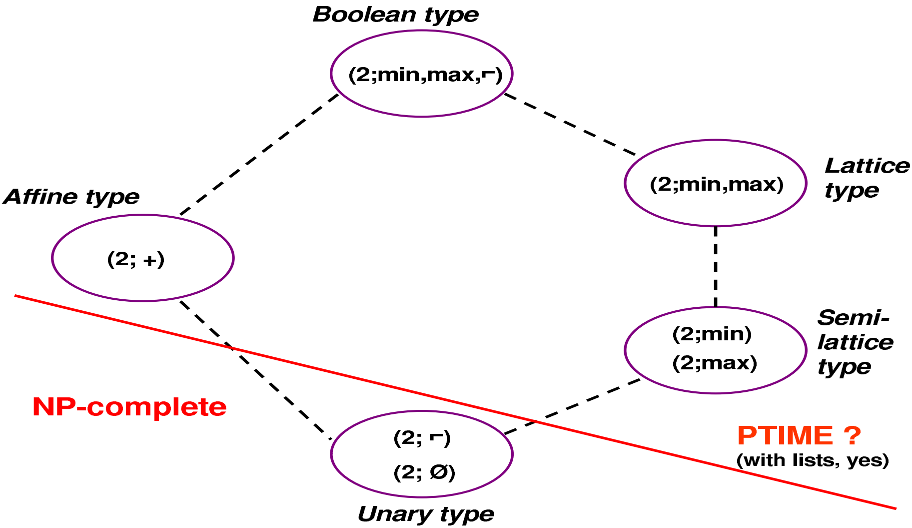
1. *(BJK) var($\mathbf{A}$) omits type 1 $\implies$ $\mathrm{CSP}(\mathbb{B})$ is in* **P**;
2. *(BL, Z) var($\mathbf{A}$) omits types 1, 2 $\Leftrightarrow$ $\neg\mathrm{CSP}(\mathbb{B})$ is in Datalog;*
3. *(BL, T) var($\mathbf{A}$) omits types 1, 2, 5 $\Leftrightarrow$ $\neg\mathrm{CSP}(\mathbb{B})$ is in Linear Datalog;*
4. *(BL, T) var($\mathbf{A}$) omits 1, 2, 4, 5 $\Leftrightarrow$ $\neg\mathrm{CSP}(\mathbb{B})$ is in Symmetric Datalog.*

- (1) = algebraic equivalent of FV dichotomy conjecture;
- (2) = analog of conjectures by Bulatov and FV.

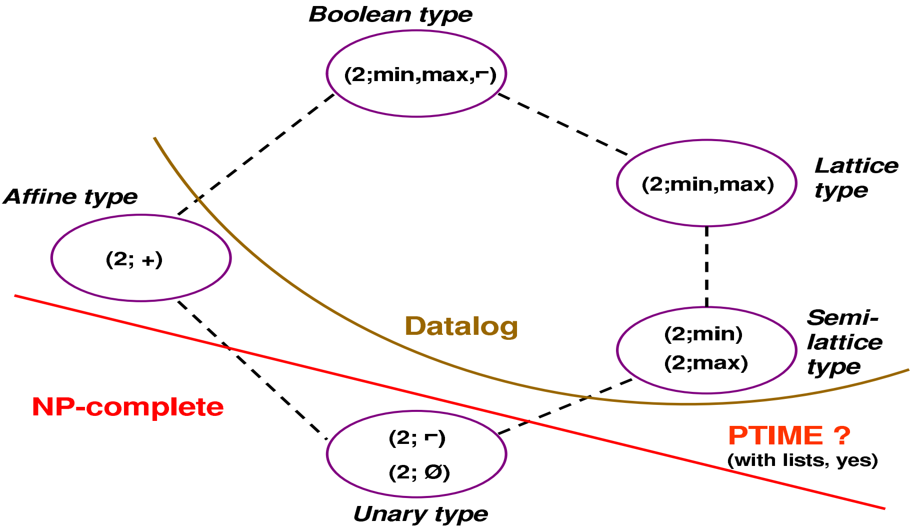
Types and Conjectures



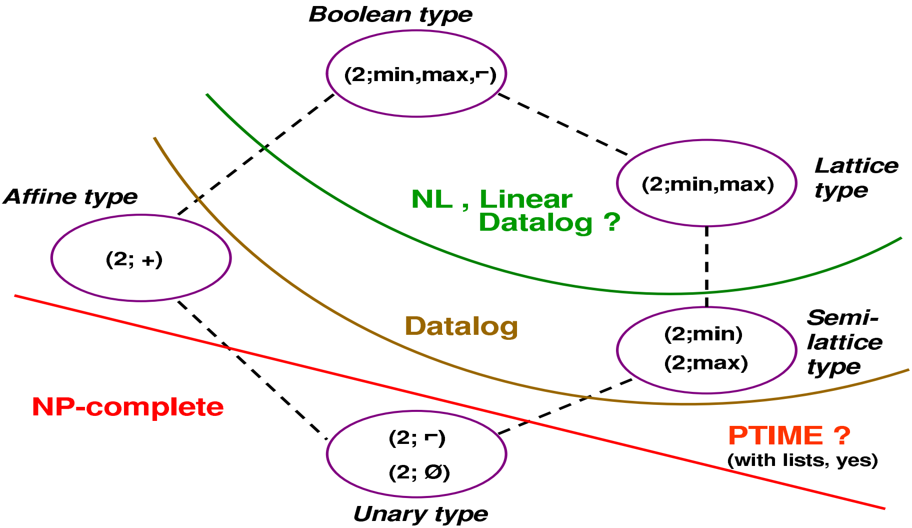
Types and Conjectures



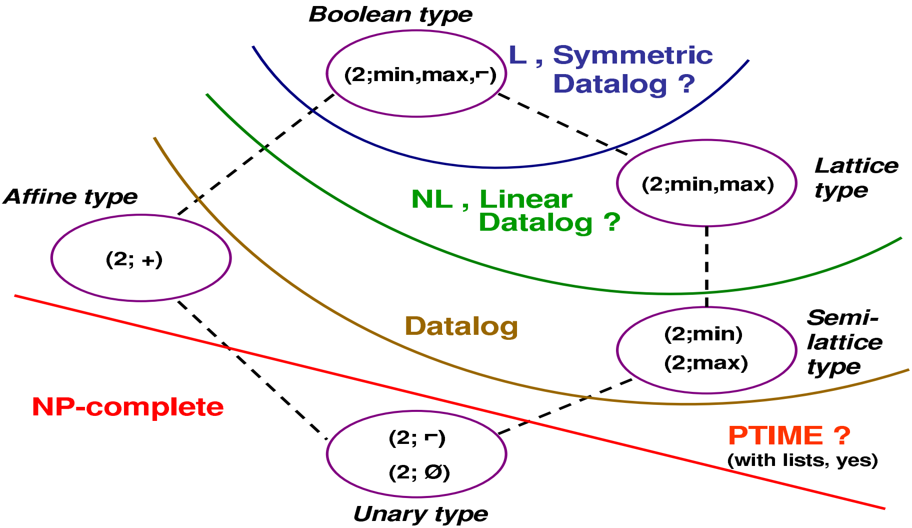
Types and Conjectures



Types and Conjectures



Types and Conjectures



Part III C: Plateau

CSPs in Datalog

$\neg\mathrm{CSP}(\mathbb{B})$ is in (l, k)-*Datalog* if solvable by a Datalog program whose rules

- have at most k variables, and
- at most l variables per IDB.

Theorem (Barto, Kozik, 2009)

Let $\mathbb{B}$ be a core structure with basic relations of arity at most r.
Then TFAE:

1. $var(\mathbf{A}(\mathbb{B}))$ *omits the unary and affine types;*
2. $\neg\mathrm{CSP}(\mathbb{B})$ *is in* $(2, \max(3, r))$-*Datalog.*

Some consequences

Theorem

Let $\mathbb{B}$ be a core structure. TFAE:

1. $var(\mathbf{A}(\mathbb{B}))$ *omits the unary and affine types;*
2. $\neg\mathrm{CSP}(\mathbb{B})$ *is in Datalog;*
3. $\neg\mathrm{CSP}(\mathbb{B})$ *is in LFP;*
4. $\neg\mathrm{CSP}(\mathbb{B})$ *is solvable by poly-size monotone circuits;*
5. $\mathrm{CSP}(\mathbb{B})$ *has a robust approximation algorithm.*

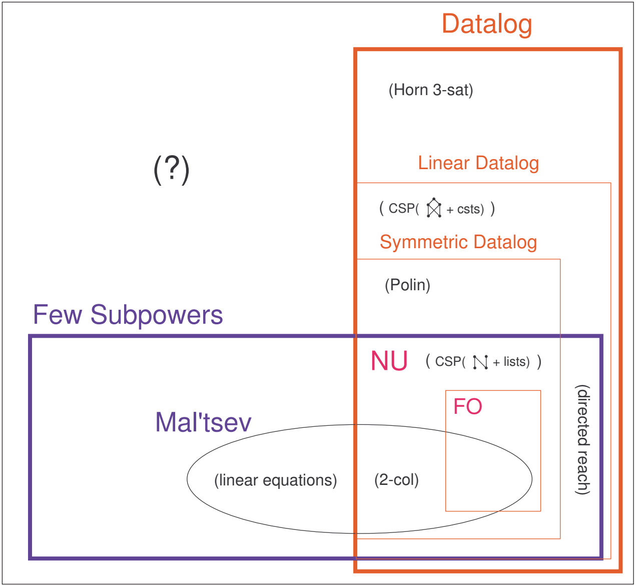
3. follows from Atserias, Bulatov, Dawar 2009;
4. follows from BL, Valeriote, Zádori 2009;
5. Barto, Kozik, 2012; as conjectured by Guruswami and Zhou.

Other Results

- Type 1 dichotomy verified for:
  - universes of size ≤ 4 (Schaefer 1978; Bulatov 2006; Márkovic 2012);
  - list-homomorphism problems (Bulatov 2003);
  - graphs (Hell, Nešetřil 1990; Bulatov 2005);
  - smooth digraphs (Barto, Kozik, Niven, 2009);
- All dichotomy conjectures verified for:
  - universe of size 2 (BL, Tesson 2007);
  - list-hom on graphs (Egri, Krokhin, BL, Tesson, 2010);
- Linear Datalog conjecture verified for
  - NU (Barto, Kozik, Willard, 2012);
- Symmetric Datalog conjecture verified for
  - Datalog + Maltsev (Dalmau, BL, 2008);

Part V: Rappelling

Tractable CSP's



Feder-Vardi (again ...)

- Feder Vardi note that polynomially generated subpowers seems to imply NOT **P**-complete;
- Idziak et al: cube terms $\implies$ CM $\implies$ omit types 1 and 5;
- So absence of type 5 seems to mean not **P**-complete;
- Feder Vardi: PGS highly parallelisable ? ($\mathbf{NC}^2$?)

From CSPs back to Algebra

Study of CSPs has fueled research in universal algebra and generated many new important concepts and results:

- Congruence Singularity (from Counting CSPs (Bulatov, Grohe))
- Cube terms, Few Subpowers; (from Dalmau's algorithm);
- Cyclic terms, Siggers terms; weak NU; (from algebraic dichotomy conjecture);
- Barto: finitely related + CD implies NU;
- Barto: finitely related + CM implies Cube term ?!! THIS JUST IN !
- McKenzie et al.: clones with cube (group) terms are finitely related;
- Barto and Kozik absorption theory
- etc.

Thank you !!

Other Trails

(added post-talk)
The algebraic approach has also been used:

- Counting CSP (full classification by Bulatov (2008);
- Quantified CSP;
- MAXCSP;
- CSPs with infinite template targets;