Reduction of CSP Dichotomy to H-Bipartite Digraphs (joint work with J. Bulin, M. Jackson, and T. Niven)

Dejan Delic
Department of Mathematics
Ryerson University
Toronto, Canada

March 15, 2012

Outline

- Fixed template constraint satisfaction problem: essentially a homomorphism problem for finite relational structures.
- We are interested in membership in the class $CSP(\mathbb{A})$, a computational problem that obviously lies in the complexity class $\mathtt{NP}$.
- Dichotomy Conjecture (Feder and Vardi): either $CSP(\mathbb{A})$ has polynomial time membership or it has $\mathtt{NP}$-complete membership problem.

Particular cases already known to exhibit the dichotomy:

- Schaefer's dichotomy for 2-element templates;
- dichotomy for undirected graph templates due to Hell and Nešetřil
- 3-element templates (Bulatov);
- digraphs with no sources and sinks (Barto, Kozik and Niven); also some special classes of oriented trees (Barto, Bulin)
- templates in which every subset is a fundamental unary relation (list homomorphism problems; Bulatov, also Barto).

- Feder and Vardi reduced the problem of proving the dichotomy conjecture to the particular case of digraph CSPs, and even to digraph CSPs whose template is a balanced digraph (a digraph on which there is a level function).
- Specifically, for every template $\mathbb{A}$ there is a balanced digraph $\mathbb{D}$ such that $CSP(\mathbb{A})$ is polynomial time equivalent to $CSP(\mathbb{D})$.
- some of the precise structure of $CSP(\mathbb{A})$ is necessarily altered in the transformation to $CSP(\mathbb{D})$.

- Algebraic approach to the CSP dichotomy conjecture: associate polynomial time algorithms to $Pol(\mathbb{A})$
- complexity of $CSP(\mathbb{A})$ is precisely (up to logspace reductions) determined by the polymorphisms of $\mathbb{A}$.
- There is some evidence that for digraphs, the algebraic structure condenses (Kazda)
- the finer structure of polymorphisms cannot in general be preserved under any translation from general CSPs to digraph CSPs.
- Maróti and Zádori: a reflexive digraph admitting Gumm polymorphisms (an extremely broad generalisation of Maltsev polymorphisms) also admits an NU polymorphism.

- Atserias (2006) revisited a construction from Feder and Vardi's original article to construct a tractable digraph CSP that is provably not solvable by the bounded width (local consistency check) algorithm.
- This construction relies heavily on finite model-theoretic machinery: quantifier preservation, cops-and-robber games (games that characterize width k), etc.

- oriented path : obtained from an undirected path by giving each edge an orientation.
- we can represent oriented paths by strings of 0's and 1's where a 0 represents a backward facing edge and a 1 represents a forward facing edge;
- for example 1001 represents the oriented path

$$\bullet \longrightarrow \bullet \longleftarrow \bullet \longleftarrow \bullet \longrightarrow \bullet.$$

- Each element a of an oriented path $\mathbb{P}$ has a level which is determined by the algebraic length of the initial segment of $\mathbb{P}$ ending at a. ($L(a)$ - level of the element a).
- balanced digraph is a digraph in which all oriented cycles have algebraic length zero.
- The vertices of such digraphs can be partitioned into levels determined by a level function L, where $L(b) = L(a) + 1$ if (a, b) is an edge.

- An oriented path is minimal if the first two edges are 11 and the last two edges are 11 and the initial and terminal vertices are the only elements on the lowest and highest levels, respectively.
- algebraic length of an oriented path $\mathbb{P}$: obtained by subtracting the number of backward facing edges from the number of forward facing edges.

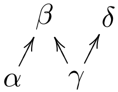
h - positive integer; $\mathbb{C}$ is h-bipartite if $\mathbb{C}$ is a balanced digraph of height h consisting of disjoint sets of vertices A and B together with a collection of (possibly isomorphic) minimal oriented paths $\mathbb{P}_1, \ldots, \mathbb{P}_k$ of algebraic length h such that

1. the set of initial vertices $\{init(\mathbb{P}_i) \mid 1 \leq i \leq k\} = A$, and
2. the set of terminal vertices $\{term(\mathbb{P}_i) \mid 1 \leq i \leq k\} = B$.
3. each element of height $l \notin \{0, h-1\}$ in $\mathbb{C}$ belongs to precisely one oriented path $\mathbb{P}_i$, that is, $(\mathbb{P}_i \cap \mathbb{P}_j) \setminus (A \cup B) = \varnothing$, for all $i \neq j$,

Main Results

- Kazda showed that a digraph admitting a Maltsev polymorphism $M(x,y,y) \approx M(y,y,x) \approx x$ (a special case of congruence modularity) necessarily has a majority polymorphism $m(x,x,y) \approx m(x,y,x) \approx m(y,x,x) \approx x$ (a ternary NU).
- Therefore, not every CSP template is pp-equivalent to a digraph template.
- Our goal is to show that every CSP template $\mathbb{A}$ is a pp-definable from a balanced digraph template $\mathbb{D}_{\mathbb{A}}$ whose CSP is polynomial time equivalent to that over $\mathbb{A}$. Moreover, every polymorphism of $\mathbb{A}$ extends to a polymorphism of $\mathbb{D}_{\mathbb{A}}$ in a way that preserves many of the most important equational properties. In this way, most of the natural equational properties that $\mathbb{A}$ has carry across to $\mathbb{D}_{\mathbb{A}}$.

The path $\mathbb{N}$



Theorem

Let $\mathbb{A}$ be a relational structure. There exists a digraph $\mathbb{D}_{\mathbb{A}}$ such that the following holds: let Σ be any linear idempotent set of identities such that each identity in Σ is either balanced or contains at most two variables. If the digraph $\mathbb{N}$ satisfies Σ, then $\mathbb{D}_{\mathbb{A}}$ satisfies Σ if and only if $\mathbb{A}$ satisfies Σ.
The digraph $\mathbb{D}_{\mathbb{A}}$ can be constructed in logspace with respect to the size of A.

Corollary

Let $\mathbb{A}$ be a CSP template. Then each of the following hold equivalently on $\mathbb{A}$ and $\mathbb{D}_{\mathbb{A}}$.

- *Taylor polymorphism or equivalently weak near-unanimity (WNU) polymorphism or equivalently cyclic polymorphism* (*conjectured to be equivalent to being tractable if $\mathbb{A}$ is a core*);
- *Polymorphisms witnessing SD*($\wedge$) (*equivalent to bounded width*);
- (*for* $k \geq 4$) *k-ary edge polymorphism* (*equivalent to few subpowers*);
- *k-ary near-unanimity polymorphism* (*equivalent to strict width*);

Corollary

(Continued)

- *totally symmetric idempotent (TSI) polymorphisms of all arities* (*equivalent to width* 1);
- *Hobby-McKenzie polymorphisms* (*equivalent to the corresponding variety satisfying a non-trivial congruence lattice identity*);
- *Gumm polymorphisms witnessing congruence modularity;*
- *Jónsson polymorphisms witnessing congruence distributivity;*
- *polymorphisms witnessing SD*$(\vee)$ (*conjectured to be equivalent to* NL*);*
- (f*or* $n \geq 3$) *polymorphisms witnessing congruence n-permutability (together with the previous item) is conjectured to be equivalent to* L*).*

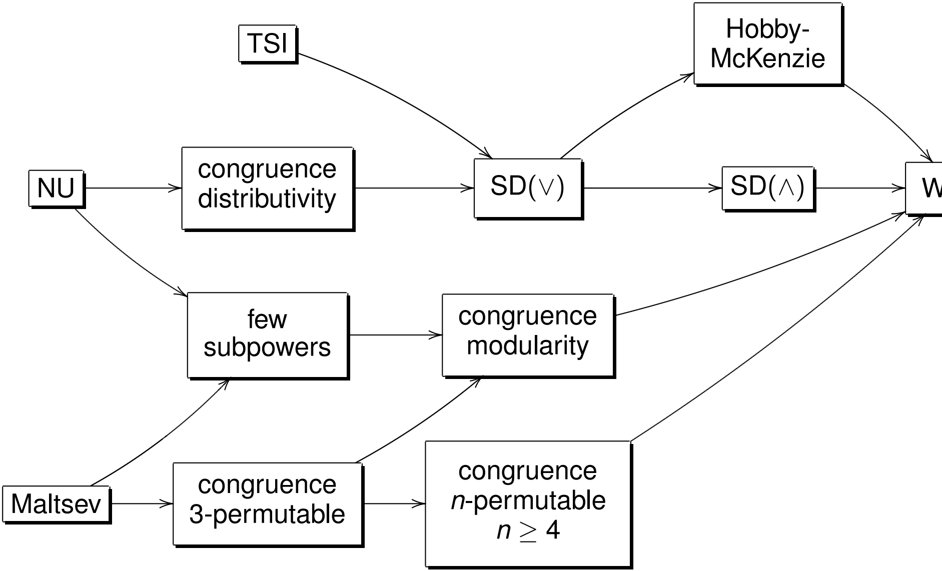


Figure: Known implications between some common Maltsev-type conditions

Lemma

The digraph $\mathbb{N}$ has a majority polymorphism, polymorphisms witnessing congruence 3-permutability and it satisfies any balanced set of identities.

Reduction to digraphs

- Can assume: $\mathbb{A}$ is not a structure with a single unary relation; also, $\mathbb{A}$ has no redundant elements.

Definition

Let $\mathbb{A} = \langle A; S_1, S_2, \ldots, S_l \rangle$ be a finite relational structure where $S_i \neq I$ has arity $k_i > 1$, for all $1 \leq i \leq l$. Define $R = S_1 \times S_2 \times \cdots \times S_l$ and view R as a k-ary relation on A where $k = \sum_{1 \leq i \leq l} k_i$.
We define a two sorted structure with k binary relations $R_1, \ldots, R_k$. Let bipartite($\mathbb{A}$) have two sorts: A and $R \subseteq A^k$. For each $i = 1, \ldots, k$ we define the relation $R_i \subseteq A \times R$ by

$$R_i := \{(x_i, (x_1, \ldots, x_n)) \mid (x_1, \ldots, x_n) \in R\}.$$

- This is a primitive positive construction, albeit a two sorted one.
- Since we assume that each element of A is involved in some relation S_i, it follows that bipartite($\mathbb{A}$) has no isolated points.
- The template $\mathbb{A}$ is primitive positive definable in bipartite($\mathbb{A}$) on the sort A by defining R as

$$R = \{(x_1, \ldots, x_k) \mid \exists y (x_1, y) \in R_1 \ \ \& \ \ \cdots \ \ \& \ \ (x_k, y) \in R_k\}.$$

Lemma

Let $\mathbb{A}$ *be a relational structure and let* Σ *be a set of identities. Then* $\mathbb{A}$ *satisfies* Σ *if and only if* $\text{bipartite}(\mathbb{A})$ *satisfies* Σ*.*

For every $e \in A \times R$, define the oriented path $\mathbb{P}_e$ (of algebraic length $2k+1$) by

$$\mathbb{P}_e = 1 + \mathbb{P}_e^1 + 1 + \mathbb{P}_e^2 + 1 + \cdots + 1 + \mathbb{P}_e^k + 1,$$

where

$$\mathbb{P}_e^i = \begin{cases} 1 & \text{if } e \in R_i \\ 101 & \text{else.} \end{cases}$$

The digraph $\mathbb{D}_{\mathbb{A}}$ is the $(2k+1)$-bipartite digraph obtained by replacing every $e = (a, b) \in A \times R$ by the oriented path $\mathbb{P}_e$ (identifying the initial and terminal vertices of $\mathbb{P}_e$ with a and b, respectively).

Lemma

The template $\mathbb{A}$ is pp-definable in $\mathbb{D}_{\mathbb{A}}$ on the set A.

Since each $\mathbb{P}_e$ is a core (minimal paths are cores) it follows that if $\mathbb{A}$ is a core, then $\mathbb{D}_{\mathbb{A}}$ is a core. Also, if $\mathbb{D}_{\mathbb{A}}$ satisfies a set of identities Σ, then $\mathbb{A}$ satisfies Σ as well. However it is not possible to pp-define $\mathbb{D}_{\mathbb{A}}$ from $\mathbb{A}$ (as $\mathbb{A}$ may not in general be pp-equivalent to a digraph).

Lemma

CSP($\mathbb{D}_{\mathbb{A}}$) reduces in polynomial time to CSP($\mathbb{A}$).

A digraph whose CSP is solvable by few-subpowers but is not bounded width

- Atserias solved a problem by Hell, Nešetřil and Zhu by giving an example of a digraph that has tractable CSP but is not of bounded width.
- The example is given by analyzing a construction of Feder-Vardi applied to AFFINE 3-SAT, and the algorithm for solving the CSP over the resulting digraph involves application of a known reduction back to AFFINE 3-SAT (which Feder and Vardi showed was not bounded width).
- our construction is based on essentially the same problem, which produces a smaller digraph and which our methods show is solvable using the few subpowers algorithm

- We consider the 2-element group $\langle\{0,1\};+\rangle$ (with $+$ interpreted modulo 2), whose term functions are known to coincide with the polymorphisms preserving the graph of the operation $+$.
- This algebra has a Maltsev term $x_1 + x_2 + x_3$, but does not generate a congruence meet-semidistributive variety.
- The relational structure $\langle\{0,1\};\text{graph}(+)\rangle$ is also not a core, as it retracts onto the induced substructure on $\{0\}$.
- add the unary relation $\{1\}$, which is still preserved by the Maltsev term $x_1 + x_2 + x_3$, and whose polymorphisms are precisely the idempotent term functions of $\langle\{0,1\};+\rangle$. It follows that the CSP over this structure, $\mathbb{A}$ say, is solvable by the few subpowers algorithm (it has a 3-edge term), but is not of bounded width.

- we replace the relations graph($+$) and $\{1\}$ by the single relation graph($+$) $\times$ $\{1\}$, which has arity 4 and involves 4 hyperedges $\{(0,0,0,1),(0,1,1,1),(1,0,1,1),(1,1,0,1)\}$; as both $\{1\}$ and $\{0\}$ are pp-definable, $\mathbb{A}$ is a core.
- The structure $\mathbb{D}_{\mathbb{A}}$ is then a core digraph (indeed it has no nontrivial automorphisms) with 102 vertices and 104 edges, has a 4-ary edge term (so CSP($\mathbb{D}_{\mathbb{A}}$) is solvable by few subpowers), but does not have bounded width.

Tractable digraphs solvable by neither few subpowers nor bounded width

For this example, we employ the “Maltsev on top” algorithm of Miklós Maróti.

- This algorithm requires a finite core relational structure $\mathbb{B}$ whose idempotent polymorphism algebra **B** has a congruence θ with the blocks of θ generating congruence meet-semidistributive varieties, and the quotient $\mathbf{B}/\theta$ having a Maltsev term.
- Consider the semigroup $\mathbf{B}'$ obtained from $\langle\{0, 1\}; +\rangle$ by adjoining an element, 2 say, that acts as a new identity element.
- There is also a congruence θ with blocks $0\,2|1$ for which the quotient $\mathbf{B}'/\theta$ is the two element group.

- Adding the two singleton relations $\{0\}$ and $\{1\}$ to $\mathbb{B}'$ produces a core relational structure, whose polymorphisms are the idempotent term functions of $\mathbf{B}'$; let this polymorphism algebra be denoted by $\mathbf{B}$.
- The quotient $\mathbf{B}/\theta$ has a Maltsev polymorphism, and shares exactly the same term functions as the polymorphism algebra of the structure $\mathbb{A}$ from the first part. Hence the Maltsev on top algorithm applies.
- The number of vertices is fairly large.

Some Problems

- As we have seen, algebraic Dichotomy Conjecture does not collapse in any substantial way on loopless digraphs.
- Can the conjectured characterizations of L and NL can be obtained on digraphs?
- How do Maróti's results (Maltsev on top, Tree on top of Maltsev) translate from **A** to $\mathbb{D}_{\mathbb{A}}$?
- Is it true that $\neg CSP(\mathbf{A})$ is expressible in some extension of FO logic if and only if $\neg HOM(\mathbb{D}_{\mathbb{A}})$ is?