Algebraic Tractability Criteria for Infinite-Domain Constraint Satisfaction Problems

Manuel Bodirsky

Institut für Informatik
Humboldt-Universität zu Berlin

May 2007
The Constraint Satisfaction Problem

Let $\Gamma = (V; R_1, \ldots, R_l)$ be a relational structure. $V$ might be infinite! Let $\tau$ be the (finite) signature of $\Gamma$.

**CSP($\Gamma$)**

**Input:** A finite $\tau$-structure $S$.

**Question:** Is there a homomorphism from $S$ to $\Gamma$?
The Constraint Satisfaction Problem

Let \( \Gamma = (V; R_1, \ldots, R_l) \) be a relational structure. \( V \) might be infinite!
Let \( \tau \) be the (finite) signature of \( \Gamma \).

**CSP(\( \Gamma \))**

**Input:** A finite \( \tau \)-structure \( S \).

**Question:** Is there a homomorphism from \( S \) to \( \Gamma \)?

**Examples:**
CSP((\( \mathbb{N}, =, \neq \))) - is \( x \) and \( y \) disconnected wrt \( = \) whenever \( x \neq y \) is in \( S \)?
CSP((\( \mathbb{Q}, < \))) - digraph acyclicity
CSP((\( \mathbb{Q}, \{(x, y, z) \mid x < y < z \lor z < y < x\}\))) - the betweenness problem
Acyclic H-colorings

Fix digraph $H$.

**Acyclic H-coloring**: (Feder+Hell+Mohar)

**Input**: A digraph $G$

**Question**: Can we $H$-color $G$ such that each color class is acyclic?
Acyclic H-colorings

Fix digraph $H$.

**Acyclic H-coloring**: (Feder+Hell+Mohar)

**Input**: A digraph $G$

**Question**: Can we $H$-color $G$ such that each color class is acyclic?

Formulation as CSP($\Gamma$): choose $\Gamma = H[(\mathbb{Q}, <)]$. 
CSPs in Temporal Reasoning

**Betweenness:**

\[
\text{CSP}((\mathbb{Q},\{(x, y, z) \mid x < y < z \lor z < y < x\}))
\]

NP-complete (Garey+Johnson)

**Min-Ordering:**

\[
\text{CSP}((\mathbb{Q},\{(x, y, z) \mid x > y \lor x > z\}))
\]

Simple linear time algorithm
Neither Datalog nor Maltsev-like
Consistency Problem for Basic Relations:

**Input:** A relational structure \((V, DR, PO, PP)\) where \(DR, PO, PP\) are binary relations

**Question:** Can we assign non-empty regions satisfying all the constraints?

Formulation as \(\text{CSP}(\Gamma)\):
Consistency Problem for Basic Relations:

Input: A relational structure \((V, DR, PO, PP)\) where \(DR, PO, PP\) are binary relations

Question: Can we assign non-empty regions satisfying all the constraints?

Formulation as CSP\((\Gamma)\):
choose \(\Gamma = (2^X \setminus \emptyset, DR, PO, PP)\) for an infinite set \(X\)
A Fundamental Lemma

Let $\mathcal{C}$ be a set of $\tau$-structures.

**Definition 1.**

$\mathcal{C}$ is **closed under disjoint unions** if whenever $A, B \in \mathcal{C}$ then $A + B \in \mathcal{C}$.

$\mathcal{C}$ is **closed under inverse homomorphisms** if $B \in \mathcal{C}$ and $A \rightarrow^h B$ implies $A \in \mathcal{C}$.

Example: the set of all triangle-free graphs

**Observation (Feder+Vardi).**

$\mathcal{C} = \text{CSP}(\Gamma)$ for some relational structure $\Gamma$ if and only if $\mathcal{C}$ is closed under disjoint unions and inverse homomorphisms.
Examples of CSPs

**Triangle-Freeness:**

**Input:** A graph $G$

**Question:** Is $G$ triangle-free?
Examples of CSPs

Triangle-Freeness:
Input: A graph $G$
Question: Is $G$ triangle-free?

Vector-space CSP:
Input: A system of linear equations $x + y = z$ and disequations $x \neq y$
Question: Is there a $d$ and an assignment of $d$-dimensional Boolean vectors to the variables that satisfies all the constraints?
1  Infinite Domain Constraint Satisfaction Problems

2  The Universal-Algebraic Approach

3  Tractability Criteria

4  Complexity Classifications
Definition 2 (First-order Interpretation).

A $\tau$-structure $\Delta$ has a interpretation in $\Gamma$ if there is

- first-order formula $\delta(x_1, \ldots, x_d)$,
- for each $m$-ary $R \in \tau$ a first-order formula $\phi_R(x_1, \ldots, x_{md})$, and
- a surjective map $h : \delta(\Gamma^d) \to \Delta$

such that for all $a_1, \ldots, a_m \in \delta(\Gamma^d)$

$$\Delta \models R(h(a_1), \ldots, h(a_m)) \iff \Gamma \models \phi_R(a_1, \ldots, a_m).$$

Note: much more powerful than first-order definitions.

An interpretation is primitive positive if $\delta$ and the $\phi_R$ are primitive positive.

Observation .

If $\Delta$ has a pp-interpretation in $\Gamma$ then there is a polynomial-time reduction from $\text{CSP}(\Delta)$ to $\text{CSP}(\Gamma)$. 
How can we recognize whether $\Gamma$ pp-interprets $\Delta$?

**Definition 3.**

A relational structure $\Gamma$ is $\omega$-categorical iff every countable model of the first-order theory of $\Gamma$ is isomorphic to $\Gamma$.

**Example:** $(\mathbb{Q}, <)$ (Kantor)

**More:** Infinite-dimensional vector-spaces over finite fields  
The countable atomless boolean algebra  
The countably infinite random graph  
The universal homogeneous triangle-free graph

**Many more:** All Fraisse-limits of amalgamation classes of structures with finite signature are $\omega$-categorical (and homogeneous)
Every CSP we have seen in this talk so far can be formulated with an \( \omega \)-categorical \( \Gamma \)!

Get new \( \omega \)-categorical structures from old:

**Observation**.

If \( \Delta \) is first-order interpretable in an \( \omega \)-categorical structure \( \Gamma \), then \( \Delta \) is also \( \omega \)-categorical.

- Allen’s Interval Algebra and all its fragments are \( \omega \)-categorical
- All CSPs in MMSNP can be formulated with an \( \omega \)-categorical template (MB+Dalmau’06)
The Basic Galois Connection

Theorem 4 (Engeler, Ryll-Nardzewski, Svenonius).

Tfae:

- $\Gamma$ is $\omega$-categorical
- $\text{Aut}(\Gamma)$ is oligomorphic, i.e., there are finitely many orbits of $k$-tuples in $\text{Aut}(\Gamma)$, for each $k$
- all orbits of $k$-tuples in $\text{Aut}(\Gamma)$ are first-order definable

Inv-Aut form a Galois connection between structures and permutation groups:

- $\text{Inv}(\text{Aut}(\Gamma))$: expansion by all first-order definable relations
- $\text{Aut}(\text{Inv}(F))$: locally closed permutation group
A Preservation Theorem

Definition 5.
A homomorphism $f$ from $\Gamma^k$ to $\Gamma$ is called a polymorphism. We say that $f$ preserves all relations in $\Gamma$.

Example: $(x, y) \mapsto \max(x, y)$ is a polymorphism of $(\mathbb{Q}, <)$, but not of $(\mathbb{Q}, Betweenness)$

Theorem 6 (MB+Nesetril’03).
A relation $R$ has a pp definition in an $\omega$-categorical structure $\Gamma$ if and only if $R$ is preserved by all polymorphisms of $\Gamma$. 
Homomorphic Equivalence

Definition 7.
Two structures $\Gamma, \Delta$ are **homomorphically equivalent** if $\Gamma$ is homomorphic to $\Delta$ and vice versa.

Example: $H[\langle \mathbb{Q}, < \rangle]$ and $\langle \mathbb{Q}, < \rangle$ are homomorphically equivalent for any finite acyclic digraph $H$.

Observation.
Two $\omega$-categorical structures $\Gamma$ and $\Delta$ are homomorphically equivalent if and only if $\text{CSP}(\Gamma)$ equals $\text{CSP}(\Delta)$. 
Cores

Let $\Gamma$ be $\omega$-categorical.

**Definition 8.**

$\Gamma$ is called a **core** if every endomorphism of $\Gamma$ is an **embedding**.  
$\Gamma$ is called **model-complete** if every embedding of $\Gamma$ into $\Gamma$ is **elementary**.

**Theorem 9.**

Every $\omega$-categorical structure $\Gamma$ is homomorphically equivalent to a model-complete core $\Delta$. Moreover,

- $\Delta$ is unique up to isomorphism
- orbits of $k$-tuples are **primitive positive** definable in $\Delta$
- $\Delta$ is $\omega$-categorical.

Consequence: can expand cores $\Delta$ by finitely many constants without changing the complexity of $\text{CSP}(\Delta)$. 

CSPs over infinite domains (May 2007)

The Universal-Algebraic Approach
Definition 10.

The algebra $\text{Al}(\Gamma)$ of $\Gamma$
- has the same domain as $\Gamma$.
- has as functions the polymorphisms of $\Gamma$.

Observation: $\text{Al}(\Gamma)$ is a locally closed clone, i.e.,
- contains all projections,
- is closed under compositions, and
- is locally closed: if for all finite subsets $S$ of the domain there is $g \in \text{Al}(\Gamma)$ s.t. $g(a) = f(a)$ for all $a \in S^k$, then $f \in \text{Al}(\Gamma)$. 
The Pseudo-Variety of an Algebra

Definition 11.
The smallest class of algebras that contains an algebra $A$ and is closed under subalgebras, homomorphic images, and finite direct products is called the \textit{pseudo-variety} $\mathcal{V}(A)$ generated by $A$.

Let $\Gamma$ be $\omega$-categorical.

Theorem 12.
A relational structure $\Delta$ has a primitive positive interpretation in $\Gamma$ if and only if there is algebra $B$ in $\mathcal{V}(\text{Alg}(\Gamma))$ all of whose operations are polymorphism of $\Delta$. 
Let $\Gamma$ be an $\omega$-categorical model-complete core.

**Theorem 13.**
If there is an expansion $\Gamma'$ of $\Gamma$ by finitely many constants such that $\forall(Al(\Gamma'))$ contains a 2-element algebra where all operations are essentially permutations, then $\text{CSP}(\Gamma)$ is NP-hard.

All hard $\omega$-categorical CSPs satisfy this condition.

**Conjecture.**
Assuming $P \neq \text{NP}$, the opposite implication is true as well.
Outline

1. Infinite Domain Constraint Satisfaction Problems
2. The Universal-Algebraic Approach
3. Tractability Criteria
4. Complexity Classifications
Outline

1 Infinite Domain Constraint Satisfaction Problems
2 The Universal-Algebraic Approach
3 Tractability Criteria
4 Complexity Classifications
Quasi Near-unanimity Operations

A quasi near-unanimity function (qnuf) is an operation satisfying

\[ f(x, \ldots, x, x, y) = f(x, \ldots, x, y, x) = \cdots = f(y, x, \ldots, x) = f(x, \ldots, x) \]

**Theorem 14 (MB+Dalmau’06).**

An \( \omega \)-categorical model-complete core \( \Gamma \) has a \( k \)-ary qnuf if and only if \( \text{CSP}(\Gamma) \) has strict bounded width (and hence, \( \text{CSP}(\Gamma) \) is tractable).

Remark: AI community says “local \( k \)-consistency implies global consistency” if \( \text{CSP}(\Gamma) \) has strict width \( k - 1 \).

Examples:

- \((\mathbb{Q},<)\) has a majority.
- \((\mathbb{Q},\leq,\neq)\) has a 5-ary, but no nuf and no 4-ary qnuf
  (MB+Chen’07 / Koubarakis)
- \((\mathbb{N},\{(x, y, u, v) \mid x \neq y \lor u \neq v\})\) has a 5-ary, but no 4-ary qnuf.
If $\Gamma = (D; R_1, \ldots, R_l)$ is a relational structure, denote by $\Gamma^c$ the expansion of $\Gamma$ by $\neg R_1, \ldots, \neg R_l$.

**Theorem 15 (MB+Chen+Kara+vonOertzen’07).**

Suppose that
- $\Gamma$ is $\omega$-categorical and admits quantifier-elimination
- $\Delta$ is first-order definable in $\Gamma$
- $\text{CSP}(\Gamma^c)$ is tractable
- there is an isomorphism $i : \Gamma^2 \to \Gamma$, and
- $\Delta$ is preserved by $i$.

Then $\text{CSP}(\Delta)$ is tractable.
If $\Gamma = (D; R_1, \ldots, R_l)$ is a relational structure, denote by $\Gamma^c$ the expansion of $\Gamma$ by $\neg R_1, \ldots, \neg R_l$.

### Theorem 15 (MB+Chen+Kara+vonOertzen’07).

Suppose that

- $\Gamma$ is $\omega$-categorical and admits quantifier-elimination
- $\Delta$ is first-order definable in $\Gamma$
- $\text{CSP}(\Gamma^c)$ is tractable
- there is an isomorphism $i : \Gamma^2 \to \Gamma$, and
- $\Delta$ is preserved by $i$.

Then $\text{CSP}(\Delta)$ is tractable.

Idea: All relations in $\Delta$ have a quantifier-free Horn definition in $\Gamma$. 

---

CSPs over infinite domains  
May 2007
Remarks

“Let $i$ be an isomorphism between $\Gamma^2$ and $\Gamma$."

1. Has no analogon in the finite!
2. Not so rare for infinite structures
3. There is automorphism $\alpha$ of $\Gamma$ such that $i$ satisfies

$$i(x, y) = \alpha(i(y, x))$$

4. All relations with a fo-definition in $\Gamma$ that are preserved by $i$ form a maximal constraint language.
Applications

Equality Constraints:

\[ \Gamma := (\mathbb{N}, =). \]
\[ \Gamma^c = (\mathbb{N}, =, \neq) \text{ clearly tractable.} \]
\[ i: \text{ any bijection between } \mathbb{N}^2 \text{ and } \mathbb{N}. \]
\[ \Delta = (\mathbb{N}, \{(x, y, u, v) \mid x = y \rightarrow u = v\}) \text{ tractable!} \]

Horn Vector-Space Equations:

\[ \Gamma := (\mathbb{V}, \{(x, y, z) \mid x + y = z\}) \text{ the infinite-dimensional vector space over a finite field} \]
\[ \Gamma^c = (\mathbb{V}, \{(x, y, z) \mid x + y = z\}), \{(x, y, z) \mid x + y \neq z\}) \text{ is tractable essentially by Gaussian elimination} \]
\[ i: \text{ an isomorphism between } \mathbb{V}^2 \text{ and } \mathbb{V}. \]
Hence: can solve Horn equations over \( \mathbb{V}. \)
More Applications

Spatial Constraints:

\[ \Gamma := (\mathbb{B}, PP, DR) \] the countable atomless boolean algebra without zero, 
\[ PP = \{(x, y) \mid xy = y\}, \quad DR = \{(x, y) \mid (x + y)x = x\} \]

\( \Gamma^c \) tractable (Renz+Nebel: Datalog)

\( i \): an isomorphism between \( \mathbb{B}^2 \) and \( \mathbb{B} \)

\( \Delta \): the maximal tractable tractable language that appeared in Drakengren+Jonsson and Renz+Nebel.

Similar applications for

- the universal triangle-free graph,
- “partially-ordered time”,
- “set-constraints”, . . .
Outline

1. Infinite Domain Constraint Satisfaction Problems
2. The Universal-Algebraic Approach
3. Tractability Criteria
4. Complexity Classifications
1 Infinite Domain Constraint Satisfaction Problems
2 The Universal-Algebraic Approach
3 Tractability Criteria
4 Complexity Classifications
An equality constraint language is a relational structure $\Gamma$ with a first-order definition in $(\mathbb{N}, =)$. 

**Examples:**

- $\Gamma = (\mathbb{N}, \neq, =)$
- $\Gamma = (\mathbb{N}, \{x = y \lor u = v\}, \{x \neq y \lor u \neq v\})$
- $\Gamma = (\mathbb{N}, \{x = y \rightarrow u = v\})$

**Theorem 16 (MB+Kara’07).** Let $\Gamma$ be an equality constraint language. Then either $V(\text{Al}(\Gamma))$ contains a two-element algebra where all ops. are projections (and CSP($\Gamma$) is NP-complete), $\Gamma$ has a polymorphism $f$, $\alpha$ satisfying $f(x, y) = \alpha(f(y, x))$ (and CSP($\Gamma$) is in P).

Proof uses polymorphisms and a Ramsey-like argument.
Equality Constraint Languages

An equality constraint language is a relational structure \( \Gamma \) with a first-order definition in \((\mathbb{N}, =)\).

Examples:

- \( \Gamma = (\mathbb{N}, \neq, =) \)
- \( \Gamma = (\mathbb{N}, \{x = y \lor u = v\}, \{x \neq y \lor u \neq v\}) \)
- \( \Gamma = (\mathbb{N}, \{x = y \rightarrow u = v\}) \)

Theorem 16 (MB+Kara’07).

Let \( \Gamma \) be an equality constraint language. Then either

- \( \text{V}(\text{Al}(\Gamma)) \) contains a two-element algebra where all ops. are projections (and \( \text{CSP}(\Gamma) \) is NP-complete),
- \( \Gamma \) has a polymorphism \( f, \alpha \) satisfying \( f(x, y) = \alpha(f(y, x)) \) (and \( \text{CSP}(\Gamma) \) is in \( \text{P} \)).
Equality Constraint Languages

An equality constraint language is a relational structure $\Gamma$ with a first-order definition in $(\mathbb{N}, =)$.

**Examples:**

- $\Gamma = (\mathbb{N}, \neq, =)$
- $\Gamma = (\mathbb{N}, \{x = y \lor u = v\}, \{x \neq y \lor u \neq v\})$
- $\Gamma = (\mathbb{N}, \{x = y \rightarrow u = v\})$

**Theorem 16 (MB+Kara’07).**

Let $\Gamma$ be an equality constraint language. Then either

- $\mathcal{V}(\text{Al}(\Gamma))$ contains a two-element algebra where all ops. are projections (and $\text{CSP}(\Gamma)$ is NP-complete),
- $\Gamma$ has a polymorphism $f$, $\alpha$ satisfying $f(x, y) = \alpha(f(y, x))$ (and $\text{CSP}(\Gamma)$ is in P).

Proof uses polymorphisms and a Ramsey-like argument.
Conclusion

1. Allowing countable ($\omega$-categorical) templates greatly expands the scope of (non-uniform) CSPs.
2. Polymorphisms are very useful to study their complexity.
3. Many more concepts from universal algebra might have generalizations to the oligomorphic setting.