CSL 2016
Marseille, August 29, 2016

Dependence Logic vs. Constraint Satisfaction

Lauri Hella and Phokion Kolaitis

Outline of the talk

- Constraint satisfaction problems
  - Feder-Vardi conjecture
  - Monotone monadic strict NP
  - MMSNP and CSP
- Dependence logic
  - Introduction
  - Team semantics
  - Uniform k-valued dependence atoms
  - Universal monotone uniform dependence logic $\forall$-MUD[ω]
- Main results
- Conclusion

Constraint satisfaction problems

A *homomorphism* between two τ-structures $\mathfrak{A}$ and $\mathfrak{B}$ is a function $h : A \to B$ such that for every $R \in \tau$, and every $(a_1, \ldots, a_n) \in A^n$,

$$(a_1, \ldots, a_n) \in R^{\mathfrak{A}} \implies (h(a_1), \ldots, h(a_n)) \in R^{\mathfrak{B}}.$$

Every τ-structure $\mathfrak{B}$ gives rise to the following constraint satisfaction problem $\mathrm{CSP}(\mathfrak{B})$:

> *Given a τ-structure $\mathfrak{A}$, does there exist a homomorphism $h : \mathfrak{A} \to \mathfrak{B}$?*

Complexity of CSP

It is easy to see that every $\mathrm{CSP}(\mathfrak{B})$ is in NP. Moreover, there are NP-complete cases.

Examples.

- Let K_n be the complete graph with n nodes. Then $\mathrm{CSP}(K_n)$ is the n-Colorability problem which is NP-complete for $n \geq 3$ and in PTIME for $n \leq 2$.

Complexity of CSP

It is easy to see that every $\mathrm{CSP}(\mathfrak{B})$ is in NP. Moreover, there are NP-complete cases.

Examples.

- Let K_n be the complete graph with n nodes. Then $\mathrm{CSP}(K_n)$ is the n-Colorability problem which is NP-complete for $n \geq 3$ and in PTIME for $n \leq 2$.
- If $\mathfrak{B} = (\{0,1\}, R^{\mathfrak{B}})$ with $R^{\mathfrak{B}} = \{(1,0,0),(0,1,0),(0,0,1)\}$, then $\mathrm{CSP}(\mathfrak{B})$ amounts to the Positive 1-in-3 Sat problem: each 3-CNF formula φ with only positive literals is encoded as $\mathfrak{A}$, where $R^{\mathfrak{A}} = \{(x,y,z) : x \vee y \vee z \text{ is a clause in } \varphi\}$.

Feder-Vardi conjecture

All known examples of $\mathrm{CSP}(\mathfrak{B})$ that are not NP-complete are in PTIME. Feder and Vardi conjectured that this is not accidental:

Dichotomy conjecture: (Feder-Vardi 98) For every structure $\mathfrak{B}$, the problem $\mathrm{CSP}(\mathfrak{B})$ is either in PTIME or NP-complete.

Feder-Vardi conjecture

All known examples of $\mathrm{CSP}(\mathfrak{B})$ that are not NP-complete are in PTIME. Feder and Vardi conjectured that this is not accidental:

Dichotomy conjecture: (Feder-Vardi 98) For every structure $\mathfrak{B}$, the problem $\mathrm{CSP}(\mathfrak{B})$ is either in PTIME or NP-complete.

The conjecture is known to be true, if

- $\mathfrak{B}$ is a graph (Hell-Nesetril 90).
- $\mathfrak{B}$ is a Boolean structure, i.e., $|B| = 2$ (Schaefer 78).
- $\mathfrak{B}$ is a three-element structure, i.e., $|B| = 3$ (Bulatov 06).

The general case of the conjecture remains unsettled.

Monotone monadic strict NP

Since all constraint satisfaction problems are in NP, by Fagin's Theorem, they are expressible in existential second-order logic Σ^1_1.

Feder and Vardi identified a natural fragment, MMSNP, of Σ^1_1 that suffices for defining $\mathrm{CSP}(\mathfrak{B})$ for every $\mathfrak{B}$.

Monotone monadic strict NP

Since all constraint satisfaction problems are in NP, by Fagin's Theorem, they are expressible in existential second-order logic Σ^1_1.

Feder and Vardi identified a natural fragment, MMSNP, of Σ^1_1 that suffices for defining $\mathrm{CSP}(\mathfrak{B})$ for every $\mathfrak{B}$.

MMSNP is defined by the following syntactic restrictions:

- all second-order quantifiers are monadic;
- all first-order quantifiers are universal;
- no inequalities occur;
- relation symbols from the underlying vocabulary occur only negatively.

MMSNP and CSP

The model-checking problem $\mathcal{MC}(\varphi)$ of an MMSNP-formula φ is:

Given a τ-structure $\mathfrak{A}$, is the case that $\mathfrak{A} \models \varphi$?

Theorem: (Feder-Vardi 98, Kun-Nesetril 08) For every $\varphi \in$ MMSNP there is a structure $\mathfrak{B}$ such that $\mathcal{MC}(\varphi)$ is equivalent to $\mathrm{CSP}(\mathfrak{B})$ w.r.t. PTIME-reductions.

Corollary: The dichotomy conjecture for CSP holds if and only if it holds for MMSNP.

Dependence logic

Dependence logic D was introduced by Väänänen 07. It is a formalism for expressing and analyzing notions of dependence and independence in computer science and mathematics, such as

- functional dependencies in relational databases;
- independence in linear algebra;
- independence in probability theory.

The origins of D can be traced back to partially ordered quantifiers (Henkin 61) and independence-friendly logic (Hintikka-Sandu 89).

Theorem: (Väänänen 07) D has the same expressive power as existential second-order logic Σ^1_1.

Dependence logic

Thus, by Fagin's Theorem, D captures NP on finite structures. In particular, every $\mathrm{CSP}(\mathfrak{B})$ is definable in D.

Surprisingly, it turns out that there are NP-complete CSPs that can be expressed already with quantifier-free formulas of D:

Dependence logic

Thus, by Fagin's Theorem, D captures NP on finite structures. In particular, every $\mathrm{CSP}(\mathfrak{B})$ is definable in D.

Surprisingly, it turns out that there are NP-complete CSPs that can be expressed already with quantifier-free formulas of D:

Theorem: (Jarmo Kontinen 13) 3-SAT can be reduced to the model-checking problem of the formula

$$\mathrm{dep}(x; y) \vee \mathrm{dep}(u; v) \vee \mathrm{dep}(u; v).$$

This raises the question whether there is a natural fragment of D that captures exactly the class of all CSPs.

Dependence logic D: Syntax

Let τ be a relational signature. The set of $\mathrm{D}(\tau)$ -formulas is defined by the following grammar:

$$\begin{array}{rcl}\varphi & ::= & x_1 = x_2 \mid \neg\, x_1 = x_2 \mid R(x_1, \ldots, x_n) \mid \neg R(x_1, \ldots, x_n) \mid \\ & & \mathrm{dep}(x_1, \ldots, x_n; y) \mid (\varphi_1 \wedge \varphi_2) \mid (\varphi_1 \vee \varphi_2) \mid \forall x \varphi \mid \exists x \varphi,\end{array}$$

where $R \in \tau$.

Note that formulas are assumed to be in *negation normal form*: negations may occur only in front of atomic formulas.

Dependence atoms occur only positively.

Dependence logic D: Team semantics

Let $\mathfrak{A}$ be a structure with domain A.

A **team** on $\mathfrak{A}$ is a set T of assignments $s : V \to A$ for some fixed set $V = \mathrm{dom}(T)$ of variables.

Dependence logic D: Team semantics

Let $\mathfrak{A}$ be a structure with domain A.

A **team** on $\mathfrak{A}$ is a set T of assignments $s : V \to A$ for some fixed set $V = \mathrm{dom}(T)$ of variables.

Literals:

- $\mathfrak{A}, T \models \lambda \iff \mathfrak{A}, s \models \lambda$ for all $s \in T$.

Dependence logic D: Team semantics

Let $\mathfrak{A}$ be a structure with domain A.

A **team** on $\mathfrak{A}$ is a set T of assignments $s : V \to A$ for some fixed set $V = \mathrm{dom}(T)$ of variables.

Literals:

- $\mathfrak{A}, T \models \lambda \iff \mathfrak{A}, s \models \lambda$ for all $s \in T$.

Connectives:

- $\mathfrak{A}, T \models \varphi \wedge \psi \iff \mathfrak{A}, T \models \varphi$ and $\mathfrak{A}, T \models \psi$.
- $\mathfrak{A}, T \models \varphi \vee \psi \iff$ there are $T', T'' \subseteq T$ s.t. $T \cup T' = T''$, $\mathfrak{A}, T' \models \varphi$ and $\mathfrak{A}, T'' \models \psi$.

Team semantics: quantifiers

To define the semantics of quantification, we use the following notation:

- $T[A/x] = \{s[a/x] \mid s \in T, a \in A\}$.
- $T[F/x] = \{s[F(s)/x] \mid s \in T\}$ for each function $F : T \to A$.

Team semantics: quantifiers

To define the semantics of quantification, we use the following notation:

- $T[A/x] = \{s[a/x] \mid s \in T, a \in A\}$.
- $T[F/x] = \{s[F(s)/x] \mid s \in T\}$ for each function $F : T \to A$.

Quantifiers:

- $\mathfrak{A}, T \models \forall x \psi \iff \mathfrak{A}, T[A/x] \models \psi$.
- $\mathfrak{A}, T \models \exists x \psi \iff$ there is $F\colon T \to A$ s.t. $\mathfrak{A}, T[F/x] \models \psi$.

Team semantics: dependence atoms

A dependence atom $\mathrm{dep}(x_1, \ldots, x_n; y)$ states that the value of y depends functionally on the values of $\boldsymbol{x} = (x_1, \ldots, x_n)$:

Dependence atoms:

- $\mathfrak{A}, T \models \mathrm{dep}(\boldsymbol{x}; y) \iff$ there is $f : A^n \to A$ s.t. for all $s \in T$, $s(y) = f(s(x_1), \ldots, s(x_n))$.

Team semantics: dependence atoms

A **uniform** dependence atom $\mathrm{udep}(x_1, \ldots, x_n; y_1, \ldots, y_n)$ states that the values of $\boldsymbol{y} = (y_1, \ldots, y_n)$ depend functionally on the values of $\boldsymbol{x} = (x_1, \ldots, x_n)$ via a unary function:

Uniform dependence atoms:

- $\mathfrak{A}, T \models \mathrm{udep}(\boldsymbol{x}; \boldsymbol{y}) \iff$ there is $g : A \to A$ s.t. for all $s \in T$,
$$s(y_1) = g(s(x_1)), \ldots, s(y_n) = g(s(x_n)).$$

Team semantics: dependence atoms

A **uniform k-valued** dep. atom $\mathrm{udep}[k](x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_n)$ states that the values of $\alpha_1, \ldots, \alpha_n$ depend functionally on the values of $x_1, \ldots, x_n$ via a unary function:

Uniform k-valued dependence atoms:

- $\mathfrak{A}, T \models \mathrm{udep}[k](\boldsymbol{x}; \boldsymbol{\alpha}) \iff$ there is $h : A \to [k]$ s.t. $\forall s \in T$,
  $s(\alpha_1) = h(s(x_1)), \ldots, s(\alpha_n) = h(s(x_n))$.

Here $\alpha_1, \ldots, \alpha_n$ are k-valued variables that range over the set $[k] = \{1, \ldots, k\}$.

Quantifier-free monotone uniform dependence logic

The syntax of *quantifier-free monotone dependence logic with uniform k-valued dependence atoms*, QF-MUD[k]:

$$\varphi \quad ::= \quad \alpha = \underline{i} \mid \neg R(\boldsymbol{x}) \mid \mathrm{udep}[k](\boldsymbol{x}; \boldsymbol{\alpha}) \mid (\varphi_1 \wedge \varphi_2) \mid (\varphi_1 \vee \varphi_2),$$

where $i \in [k]$.

The union of QF-MUD[k] over all $k \geq 1$ is denoted by QF-MUD[ω].

Universal monotone uniform dependence logic

The syntax of *universal monotone dependence logic with uniform k-valued dependence atoms*, $\forall$-MUD[k]:

$$\varphi \quad ::= \quad \psi \mid \forall x \varphi \mid \forall \alpha \varphi,$$

where $\psi \in$ QF-MUD[k].

The union of $\forall$-MUD[k] over all $k \geq 1$ is denoted by $\forall$-MUD[ω].

Universal monotone uniform dependence logic

The syntax of *universal monotone dependence logic with uniform k-valued dependence atoms*, $\forall$-MUD[k]:

$$\varphi \quad ::= \quad \psi \mid \forall x \varphi \mid \forall \alpha \varphi,$$

where $\psi \in$ QF-MUD[k].

The union of $\forall$-MUD[k] over all $k \geq 1$ is denoted by $\forall$-MUD[ω].

Analogously to MMSNP, the logics QF-MUD[k] and $\forall$-MUD[k] admit no inequalities and only negative occurrences of $R \in \tau$.

Main results

We prove first that any $\mathrm{CSP}(\mathfrak{B})$ is definable in $\forall\text{-}\mathrm{MUD}[\omega]$, assuming that $\mathfrak{B}$ has only one relation. This suffices, since every $\mathrm{CSP}(\mathfrak{C})$ is PTIME-equivalent to such $\mathrm{CSP}(\mathfrak{B})$.

Main results

We prove first that any $\mathrm{CSP}(\mathfrak{B})$ is definable in $\forall$-MUD$[\omega]$, assuming that $\mathfrak{B}$ has only one relation. This suffices, since every $\mathrm{CSP}(\mathfrak{C})$ is PTIME-equivalent to such $\mathrm{CSP}(\mathfrak{B})$.

Theorem A: ($\forall$-MUD$[\omega]$ captures CSP)
Let $\tau = \{R\}$, and let $\mathfrak{B}$ be a τ-structure with $|B| = k$. There is a sentence $\varphi_{\mathfrak{B}} \in \forall$-MUD$[k]$ such that

$$\mathfrak{A} \models \varphi_{\mathfrak{B}} \iff \mathfrak{A} \in \mathrm{CSP}(\mathfrak{B})$$

holds for all τ-structures $\mathfrak{A}$.

Main results

On the other hand, we prove that MMSNP is at least as expressive as $\forall$-MUD[ω].

Theorem B: ($\forall$-MUD[ω] is contained in MMSNP)
Let $\varphi \in \forall\text{-MUD}[\omega]$ be a sentence. There is a sentence $\varphi^* \in$ MMSNP such that

$$\mathfrak{A} \models \varphi \iff \mathfrak{A} \models \varphi^*$$

holds for all τ-structures $\mathfrak{A}$.

Main results

On the other hand, we prove that MMSNP is at least as expressive as $\forall$-MUD$[\omega]$.

Theorem B: ($\forall$-MUD$[\omega]$ is contained in MMSNP)
Let $\varphi \in \forall\text{-MUD}[\omega]$ be a sentence. There is a sentence $\varphi^* \in$ MMSNP such that

$$\mathfrak{A} \models \varphi \iff \mathfrak{A} \models \varphi^*$$

holds for all τ-structures $\mathfrak{A}$.

Corollary: The dichotomy conjecture for CSP holds if and only if it holds for $\forall$-MUD$[\omega]$.

Expressing CSP in universal monotone uniform DL

To prove Theorem A, it suffices to find a formula $\psi_{\mathfrak{B}} \in \text{QF-MUD}[k]$ such that

$$\mathfrak{A}, F \models \psi_{\mathfrak{B}} \iff \mathfrak{A} \in \mathrm{CSP}(\mathfrak{B}),$$

where F is the *full team* consisting of all assignments $s : \{x_1, \dots, x_n, \alpha_1, \dots, \alpha_n\} \to A \cup [k]$.

This is because $\mathfrak{A}, F \models \psi_{\mathfrak{B}} \iff \mathfrak{A} \models \varphi_{\mathfrak{B}}$, where $\varphi_{\mathfrak{B}}$ is the sentence $\forall \boldsymbol{x} \forall \boldsymbol{\alpha}\, \psi_{\mathfrak{B}}$.

Expressing CSP in universal monotone uniform DL

Observe next that if $\mathfrak{A}, T \models \mathrm{udep}[k](\boldsymbol{x}, \boldsymbol{\alpha})$, then there is a homomorphism $h : (A, R_{T,\boldsymbol{x}}) \to ([k], R_{T,\boldsymbol{\alpha}})$.

Thus, if $R^{\mathfrak{A}} \subseteq R_{T,\boldsymbol{x}}$ and $R_{T,\boldsymbol{\alpha}} \subseteq R^{\mathfrak{B}}$, then $\mathfrak{A} \in \mathrm{CSP}(\mathfrak{B})$.

(Here $R_{T,\boldsymbol{x}}$ is the relation $\{s(\boldsymbol{x}) : s \in T\}$, and similarly for $R_{T,\boldsymbol{\alpha}}$.)

Expressing CSP in universal monotone uniform DL

Observe next that if $\mathfrak{A}, T \models \mathrm{udep}[k](\boldsymbol{x}, \boldsymbol{\alpha})$, then there is a homomorphism $h : (A, R_{T,\boldsymbol{x}}) \to ([k], R_{T,\boldsymbol{\alpha}})$.

Thus, if $R^{\mathfrak{A}} \subseteq R_{T,\boldsymbol{x}}$ and $R_{T,\boldsymbol{\alpha}} \subseteq R^{\mathfrak{B}}$, then $\mathfrak{A} \in \mathrm{CSP}(\mathfrak{B})$.

(Here $R_{T,\boldsymbol{x}}$ is the relation $\{s(\boldsymbol{x}) : s \in T\}$, and similarly for $R_{T,\boldsymbol{\alpha}}$.)

The idea of the proof is to build $\psi_{\mathfrak{B}}$ (using disjunctions) in such a way that if $\mathfrak{A}, F \models \psi_{\mathfrak{B}}$, then there is a subteam T of F satisfying the conditions above.

Universal monotone uniform DL and MMSNP

In the proof of Theorem B we define inductively a translation of QF-MUD[k] to the extension of MMSNP with k-valued variables.

The translation of each formula $\psi \in$ QF-MUD[k] is of the form $\exists P_1 \ldots \exists P_{k\ell} \forall \boldsymbol{x} \forall \boldsymbol{y} \forall \boldsymbol{\alpha} (R(\boldsymbol{x}\boldsymbol{\alpha}) \to \psi^+)$, where ψ^+ is quantifier free.

Universal monotone uniform DL and MMSNP

In the proof of Theorem B we define inductively a translation of QF-MUD[k] to the extension of MMSNP with k-valued variables.

The translation of each formula $\psi \in$ QF-MUD[k] is of the form $\exists P_1 \ldots \exists P_{k\ell} \forall \boldsymbol{x} \forall \boldsymbol{y} \forall \boldsymbol{\alpha}(R(\boldsymbol{x\alpha}) \to \psi^+)$, where ψ^+ is quantifier free.

Note that here R is an extra relation symbol, whose interpretation is given by the team on which ψ is evaluated: we prove that

$$\mathfrak{A}, T \models \psi \iff (\mathfrak{A}, R_{T,\boldsymbol{x\alpha}}) \models \exists \mathbf{P} \forall \boldsymbol{x} \forall \boldsymbol{y} \forall \boldsymbol{\alpha}(R(\boldsymbol{x\alpha}) \to \psi^+).$$

This extends in a straightforward way to sentences of $\forall$-MUD[k]:

$$\mathfrak{A}, T \models \forall \boldsymbol{x} \forall \boldsymbol{\alpha} \psi \iff \mathfrak{A} \models \exists \mathbf{P} \forall \boldsymbol{x} \forall \boldsymbol{y} \forall \boldsymbol{\alpha} \psi^+.$$

Theorem B follows from this, as the k-valued variables can be easily eliminated from sentences of MMSNP.

Conclusion

Our main results settle the complexity of model-checking for $\forall$-MUD$[\omega]$: for each sentence $\varphi \in \forall$-MUD$[\omega]$, the problem $\mathcal{MC}(\varphi)$ is PTIME-equivalent to some $\mathrm{CSP}(\mathfrak{B})$, and vice versa.

However, the exact expressive power of $\forall$-MUD$[\omega]$ is not clear yet:

- Is $\mathrm{CSP}(\mathfrak{B})$ definable in $\forall$-MUD$[\omega]$ for every $\mathfrak{B}$?
- Is $\forall$-MUD$[\omega]$ a proper fragment of MMSNP?

Conclusion

Our main results settle the complexity of model-checking for $\forall$-MUD$[\omega]$: for each sentence $\varphi \in \forall$-MUD$[\omega]$, the problem $\mathcal{MC}(\varphi)$ is PTIME-equivalent to some $\mathrm{CSP}(\mathfrak{B})$, and vice versa.

However, the exact expressive power of $\forall$-MUD$[\omega]$ is not clear yet:

- Is $\mathrm{CSP}(\mathfrak{B})$ definable in $\forall$-MUD$[\omega]$ for every $\mathfrak{B}$?
- Is $\forall$-MUD$[\omega]$ a proper fragment of MMSNP?

Another natural question concerns the relationship between the uniform and the ordinary dependence atoms:

- Is $\mathrm{udep}[k](\boldsymbol{x}; \boldsymbol{\alpha})$ definable in the universal fragment of D?

Since D captures Σ^1_1, $\mathrm{udep}[k](\boldsymbol{x}; \boldsymbol{\alpha})$ is definable in full D, but its definition seems to require existential quantification.

Thanks for your attention!