Min CSP on Four Elements: Moving Beyond Submodularity

Johan Thapper

LIX, École Polytechnique

joint work with
Peter Jonsson (IDA, Linköping University) and Fredrik Kuivinen

Workshop on Algebra and CSPs
The Fields Institute, Toronto, August 2011

Outline

Outline

MAX CSP

Definition (MAX CSP(Γ))

Let Γ be a finite constraint language over a finite domain D.

Instance: A CSP(Γ)-instance I.

Goal: Find an assignment to the variables in I which maximises the number of satisfied constraints.

Max CSP

Definition (Max CSP(Γ))

Let Γ be a finite constraint language over a finite domain D.

Instance: A CSP(Γ)-instance I.

Goal: Find an assignment to the variables in I which maximises the number of satisfied constraints.

Some (NP-hard) examples:

- Max cut ($\Gamma = \{\mathrm{XOR}\}$)
- Max k-cut ($\Gamma = \{\neq_k\}$)
- Max k-SAT ($\Gamma = \{\mathrm{OR}_{i,j} \mid 0 \leq j \leq i \leq k\}$)

where $\mathrm{XOR} := \{(0,1),(1,0)\}$, $\neq_k := \{(a,b) \in D^2 \mid a \neq b\}$ $(k = |D|)$, and $\mathrm{OR}_{i,j} := \{(x_1,\ldots,x_i) \in \{0,1\}^i \mid \neg x_1 \vee \cdots \vee \neg x_j \vee x_{j+1} \vee \cdots \vee x_i\}$.

PO versus NP-hard

- The algorithms of Raghavendra (2008), and Raghavendra and Steurer (2009) give optimal approximation ratios for $\textsc{Max}\ \mathrm{CSP}(\Gamma)$ under the assumption of the **Unique Games Conjecture** (UGC).

PO versus NP-hard

- The algorithms of Raghavendra (2008), and Raghavendra and Steurer (2009) give optimal approximation ratios for $\textsc{Max}\ \mathrm{CSP}(\Gamma)$ under the assumption of the **Unique Games Conjecture** (UGC).
- For problems in PO these techniques give a PTAS, but it is not clear how to determine in general when this is the case.

PO versus NP-hard

- The algorithms of Raghavendra (2008), and Raghavendra and Steurer (2009) give optimal approximation ratios for $\textsc{Max}\ \mathrm{CSP}(\Gamma)$ under the assumption of the **Unique Games Conjecture** (UGC).
- For problems in PO these techniques give a PTAS, but it is not clear how to determine in general when this is the case.
- We address the problem of tracing the boundary between problems inside and outside of PO.

PO versus NP-hard

- The algorithms of Raghavendra (2008), and Raghavendra and Steurer (2009) give optimal approximation ratios for $\text{Max CSP}(\Gamma)$ under the assumption of the **Unique Games Conjecture** (UGC).
- For problems in PO these techniques give a PTAS, but it is not clear how to determine in general when this is the case.
- We address the problem of tracing the boundary between problems inside and outside of PO.
- We find a dichotomy for constraint languages with domain size 4 between problems in PO and NP-hard problems.

PO versus NP-hard

- The algorithms of Raghavendra (2008), and Raghavendra and Steurer (2009) give optimal approximation ratios for $\text{Max CSP}(\Gamma)$ under the assumption of the **Unique Games Conjecture** (UGC).
- For problems in PO these techniques give a PTAS, but it is not clear how to determine in general when this is the case.
- We address the problem of tracing the boundary between problems inside and outside of PO.
- We find a dichotomy for constraint languages with domain size 4 between problems in PO and NP-hard problems.
- We identify a new type of tractable problems which has not previously shown up in classifications of Max CSP.

Max CSP using $\{0, 1\}$-functions

We represent a k-ary relation $R \in \Gamma$ by its characteristic function $f : D^k \to \{0, 1\}$, with $f(\mathbf{x}) = 1$ iff $\mathbf{x} \in R$.

Definition ((Weighted) Max CSP(Γ))

Let Γ be a set of $\{0, 1\}$-valued functions over D.

Instance: A formal sum $\sum_{i=1}^{n} w_i f_i(\mathbf{x_i})$, where $w_i \in \mathbb{Q}_{\geq 0}$, f_i is a k_i-ary cost function in Γ, and $\mathbf{x_i} \in V^{k_i}$.

Solution: A function $\sigma : V \to D$.

Measure: $\sum_{i=1}^{n} w_i f_i(\sigma(\mathbf{x_i}))$, where σ is applied componentwise.

MIN CSP

In order to study MAX CSP in the VCSP-framework, we choose to work with MIN CSP instead. The reason for this is to keep certain terminology consistent (multimorphisms, submodularity, etc.).

Observation

Let $\Gamma^c = \{1 - f \mid f \in \Gamma\}$. *The problems* MAX CSP$(\Gamma)$ *and* MIN CSP(Γ^c) *are polynomial-time equivalent.*

Note that the two problems may differ with respect to approximability.

Known classification results

Full classifications of $\textsc{Min CSP}(\Gamma)$ exist for the following cases:

- 2-element domains; Creignou (1995)
- 3-element domains; Jonsson, Klasson, and Krokhin (2006)
- Γ containing a single function; Jonsson and Krokhin (2007)
- Γ containing all unary functions; Deineko, Jonsson, Klasson, and Krokhin (2008)

In each of these cases, provided that Γ is a core, $\textsc{Min CSP}(\Gamma)$ is tractable if and only if Γ **is submodular with respect to some chain on** D.

Known classification results

Full classifications of $\textsc{Min CSP}(\Gamma)$ exist for the following cases:

- 2-element domains; Creignou (1995)
- 3-element domains; Jonsson, Klasson, and Krokhin (2006)
- Γ containing a single function; Jonsson and Krokhin (2007)
- Γ containing all unary functions; Deineko, Jonsson, Klasson, and Krokhin (2008)

In each of these cases, provided that Γ is a core, $\textsc{Min CSP}(\Gamma)$ is tractable if and only if Γ **is submodular with respect to some chain on** D.

$\textsc{Min CSP}(\Gamma)$ is also tractable when

- Γ is submodular with respect to any distributive lattice.
- Γ is submodular with respect to certain non-distributive lattices; Krokhin and Larose (2007), Kuivinen (2009)

Beyond submodularity

Given the previously known evidence, a tentative conjecture has been that, for a core Γ,

- $\textsc{Min}\ \mathrm{CSP}(\Gamma)$ is tractable when Γ is submodular with respect to some lattice, and
- this is the only source of tractability for $\textsc{Min}\ \mathrm{CSP}(\Gamma)$.

Beyond submodularity

Given the previously known evidence, a tentative conjecture has been that, for a core Γ,

- $\textsc{Min CSP}(\Gamma)$ is tractable when Γ is submodular with respect to some lattice, and
- this is the only source of tractability for $\textsc{Min CSP}(\Gamma)$.

We show that the second part is false.

Counterexample

Let $D = \{a, b, c, d\}$, and let $\Gamma = \{\{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}, R\}$, where $R = \{(x, y) \mid x = b \vee y = c\}$. Then Γ is a core which is not submodular with respect to any lattice on D, yet $\textsc{Min CSP}(\Gamma)$ is tractable.

VCSP without mixed cost functions

It will be useful to allow "crisp" constraints in some of the reductions. There will be no "mixed" constraints, i.e., cost functions taking both a non-zero finite and an infinite value.

Definition

Let Γ be a set of finite-valued cost functions on a domain D, and Δ be a set of relations on D. $\mathrm{VCSP}(\Gamma, \Delta)$ is the following minimisation problem:

Instance: A formal sum $\sum_{i=1}^{n} w_i f_i(\mathbf{x_i})$, and a finite set of constraint applications $\{(\mathbf{y_j}; R_j)\}$, where $f_i \in \Gamma$, $R_j \in \Delta$, and $\mathbf{x_i}, \mathbf{y_j}$ are matching lists of variables from V.

Solution: A function $\sigma : V \to D$ such that $\sigma(\mathbf{y_j}) \in R_j$ for all j.

Measure: $\sum_{i=1}^{n} w_i f_i(\sigma(\mathbf{x_i}))$.

We write $\mathrm{VCSP}(\Gamma)$ when Δ is empty and $\textsc{Min}\ \mathrm{CSP}(\Gamma, \Delta)$ when Γ consists of $\{0, 1\}$-valued constraints only.

Outline

Multimorphisms

Cohen, Cooper, Jeavons, and Krokhin (2006) introduced the first in a number of generalisations of polymorphisms for the purpose of investigating the VCSP.

Definition

A pair of functions $f, g : D^2 \to D$ is a (binary) **multimorphism** of a k-ary cost function $h : D^k \to \mathbb{Q}_{\geq 0}$ if, for all tuples $\mathbf{x}, \mathbf{y} \in D^k$,

$$h(\mathbf{x}) + h(\mathbf{y}) \geq h(f(\mathbf{x}, \mathbf{y})) + h(g(\mathbf{x}, \mathbf{y})),$$

where f and g are applied component-wise. It is a multimorphism of (Γ, Δ) if it is a multimorphism of each function in $\Gamma \cup \Delta$.

Multimorphisms; submodular cost functions

Example

A cost function $h : D^k \to \mathbb{Q}$ is **submodular** with respect to a lattice $L = (D; \wedge, \vee)$ if

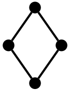
$$f(\mathbf{a}) + f(\mathbf{b}) \geq f(\mathbf{a} \wedge \mathbf{b}) + f(\mathbf{a} \vee \mathbf{b}),$$

for all $\mathbf{a}, \mathbf{b} \in D^k$.

In this case, $(\wedge, \vee)$ is a multimorphism of h.

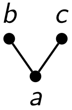
Theorem (Schrijver (2000); Iwata, Fleischer, and Fujishige (2001))

If Γ is submodular w.r.t. a distributive lattice, then $\mathrm{VCSP}(\Gamma)$ is tractable.

Examples include chains (total orders) and products of chains, e.g., 

Bisubmodular cost functions

Let $D = \{a, b, c\}$, and let $\sqcap, \sqcup : D \to D$ be the $\wedge$ and $\vee$ of the poset



on pairs for which they are defined, and $\sqcup(b, c) = \sqcup(c, b) = a$.

$\sqcap$	a	b	c
a	a	a	a
b	a	b	a
c	a	a	c

$\sqcup$	a	b	c
a	a	b	c
b	b	b	a
c	c	a	c

Theorem (Fujishige and Iwata (2006))

If Γ has the multimorphism $(\sqcap, \sqcup)$, then $\mathrm{VCSP}(\Gamma)$ is tractable.

Congruences

Definition

Let $(D; f, g)$ be an algebra with two binary operations, f and g.
A **congruence** of $(D; f, g)$ is an equivalence relation θ such that

$$(x_1, x_2), (y_1, y_2) \in \theta \implies (f(x_1, y_1), f(x_2, y_2)), (g(x_1, y_1), g(x_2, y_2)) \in \theta,$$

for all $x_1, x_2, y_1, y_2 \in D$.

- D/θ denotes the set of equivalence classes in θ.
- $x[\theta] \in D/\theta$ denotes the equivalence class of $x \in D$.
- The following are well-defined operations on D/θ:

$$f/\theta(x[\theta], y[\theta]) := f(x, y)[\theta] \qquad g/\theta(x[\theta], y[\theta]) := g(x, y)[\theta].$$

Mal'tsev products

Definition

Let $\mathbf{V}$ and $\mathbf{W}$ be classes of algebras $(D'; f', g')$, with f' and g' binary operations. The **Mal'tsev product**, $\mathbf{V} \circ \mathbf{W}$, consists of all algebras $(D; f, g)$ such that there is a congruence θ where every class supports a subalgebra, and

1. $(x[\theta], f|_{x[\theta]}, g|_{x[\theta]}) \in \mathbf{V}$ for each $x[\theta] \in D/\theta$; and
2. $(D/\theta, f/\theta, g/\theta) \in \mathbf{W}$.

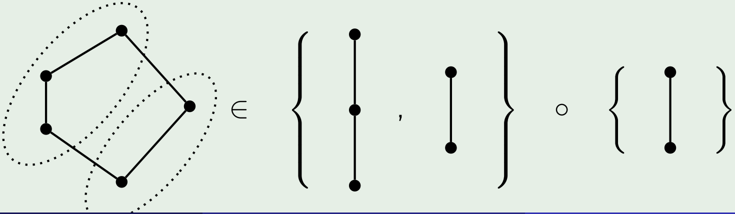
Mal'tsev products

Definition

Let **V** and **W** be classes of algebras $(D'; f', g')$, with f' and g' binary operations. The **Mal'tsev product**, $\mathbf{V} \circ \mathbf{W}$, consists of all algebras $(D; f, g)$ such that there is a congruence θ where every class supports a subalgebra, and

1. $(x[\theta], f|_{x[\theta]}, g|_{x[\theta]}) \in \mathbf{V}$ for each $x[\theta] \in D/\theta$; and
2. $(D/\theta, f/\theta, g/\theta) \in \mathbf{W}$.

Example (The pentagon lattice)



MFM

Definition (Multimorphism Function Minimisation)

Let X be a finite set of triples $(D_i; f_i, g_i)$, where D_i is a finite set and $f_i, g_i : D_i^2 \to D_i$. $\mathrm{MFM}(X)$ is the minimisation problem:

Instance: A positive integer n, a function $j : \{1, \ldots, n\} \to \{1, \ldots, |X|\}$, and a function $h : \mathcal{D} \to \mathbb{Z}$, where $\mathcal{D} = \prod_{i=1}^{n} D_{j(i)}$. For all $\mathbf{x}, \mathbf{y} \in \mathcal{D}$,

$$h(\mathbf{x}) + h(\mathbf{y}) \geq h(f_{j(1)}(x_1, y_1), f_{j(2)}(x_2, y_2), \ldots, f_{j(n)}(x_n, y_n)) + h(g_{j(1)}(x_1, y_1), g_{j(2)}(x_2, y_2), \ldots, g_{j(n)}(x_n, y_n)).$$

The function h is assumed be supplied as an oracle.

Solution: A tuple $\mathbf{x} \in \mathcal{D}$.

Measure: The value of $h(\mathbf{x})$.

MFM and Mal'tsev products

- MFM(X) is said to be **oracle-tractable** if it can be solved in polynomial time in the number of variables of the input function h.
- If Γ has the multimorphism (f, g), and MFM$(\{(D; f, g)\})$ is oracle-tractable, then VCSP(Γ) is tractable.

Theorem (cf. Krokhin and Larose (2007), Theorem 4.3)

Suppose that $\mathbf{V}$ *and* $\mathbf{W}$ *are finite sets of triples* $(D; f, g)$ *such that MFM(*$\mathbf{V}$*) and MFM(*$\mathbf{W}$*) are oracle-tractable. Then MFM(*$\mathbf{V} \circ \mathbf{W}$*) is also oracle-tractable.*

MFM and Mal'tsev products

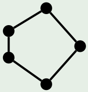
- MFM(X) is said to be **oracle-tractable** if it can be solved in polynomial time in the number of variables of the input function h.
- If Γ has the multimorphism (f, g), and MFM$(\{(D; f, g)\})$ is oracle-tractable, then VCSP(Γ) is tractable.

Theorem (cf. Krokhin and Larose (2007), Theorem 4.3)

Suppose that $\mathbf{V}$ *and* $\mathbf{W}$ *are finite sets of triples* $(D; f, g)$ *such that MFM(*$\mathbf{V}$*) and MFM(*$\mathbf{W}$*) are oracle-tractable. Then MFM(*$\mathbf{V} \circ \mathbf{W}$*) is also oracle-tractable.*

Example

MFM({  ,  }) and MFM({  }) are oracle-tractable, so

MFM({  }) is also oracle-tractable.

1-defect chains

Definition

Let $(D; <)$ be a chain, and let $b, c \in D$ be two distinct elements. Assume that $f, g : D^2 \to D$ are two commutative operations such that:

- $\{x, y\} \neq \{b, c\} \implies f(x, y) = \min_<(x, y), g(x, y) = \max_<(x, y)$;
- $\{f(b, c), g(b, c)\} \cap \{b, c\} = \emptyset$; and
- $f(b, c) < g(b, c)$.

We call $(D; f, g)$ a **1-defect chain**, and we say that (f, g) is a **1-defect chain multimorphism**.

1-defect chains

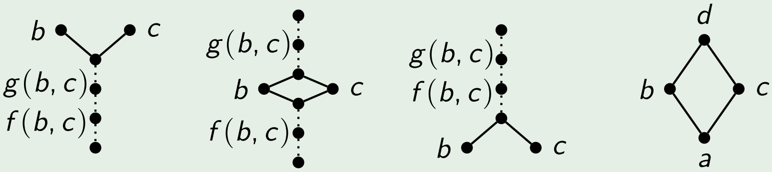
Definition

Let $(D;<)$ be a chain, and let $b, c \in D$ be two distinct elements. Assume that $f, g : D^2 \to D$ are two commutative operations such that:

- $\{x, y\} \neq \{b, c\} \implies f(x, y) = \min_<(x, y), g(x, y) = \max_<(x, y)$;
- $\{f(b, c), g(b, c)\} \cap \{b, c\} = \emptyset$; and
- $f(b, c) < g(b, c)$.

We call $(D; f, g)$ a **1-defect chain**, and we say that (f, g) is a **1-defect chain multimorphism**.

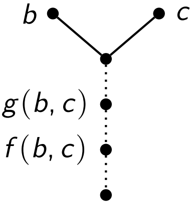
Example



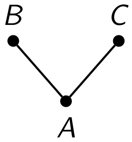
Tractability for 1-defect chain multimorphisms

Proposition

If Γ has a 1-defect chain multimorphism, then VCSP(Γ) is tractable.



$\Longrightarrow$



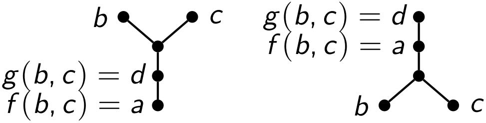
$= f/\theta(B, C) = g/\theta(B, C)$

Let $D/\theta = \{A, B, C\}$ with $A = D \setminus \{b, c\}$, $B = \{b\}$, $C = \{c\}$.

- θ is a congruence;
- $(A; f|_A, g|_A)$, $(B; f|_B, g|_B)$, and $(C; f|_C, g|_C)$ are chains; and
- $(D/\theta; f/\theta, g/\theta)$ is isomorphic to $(\{a, b, c\}; \sqcap, \sqcup)$.

Example

- $D = \{a, b, c, d\}$
- $$\Gamma = \left\{ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$
- Γ is a core which is not submodular with respect to any lattice on D, but it has the following 1-defect chain multimorphisms.



Outline

Weighted pp-definitions

Let I be an instance of $\mathrm{VCSP}(\Gamma, \Delta)$ with variables $(v_1, \ldots, v_n)$.

$$\mathrm{Optsol}(I) := \{(\sigma(v_1), \ldots, \sigma(v_n)) \mid \sigma \text{ is an optimal solution to } I\}$$

A relation has a **weighted pp-definition** over (Γ, Δ) if it is equal to $\pi_{\mathbf{x}}\mathrm{Optsol}(I)$ for some I and sequence $\mathbf{x}$ of variables from I.

Let $\langle \Gamma, \Delta \rangle_w$ denote the set of all such relations.

Proposition

Let $\Delta' \subseteq \langle \Gamma, \Delta \rangle_w$ be a finite subset. Then $\mathrm{VCSP}(\Gamma, \Delta')$ is polynomial-time reducible to $\mathrm{VCSP}(\Gamma, \Delta)$.

Endomorphisms and cores for $\textsc{Min CSP}$

Let Γ be a finite set of $\{0,1\}$-valued cost functions.

- An operation $f : D \to D$ is an **endomorphism** of Γ if

$$h(\mathbf{a}) = 0 \implies h(f(\mathbf{a})) = 0.$$

- An **automorphism** is a surjective endomorphism.
- The set of endomorphisms (automorphisms) of Γ is denoted by $\mathrm{End}(\Gamma)$ $(\mathrm{Aut}(\Gamma))$.
- Γ is called a **core** if all of its endomorphisms are automorphisms.

Lemma

Let $f \in \mathrm{End}(\Gamma)$, $D' = f(D)$, and $\Gamma' = \{h|_{D'} \mid h \in \Gamma\}$. Then $\textsc{Min CSP}(\Gamma)$ and $\textsc{Min CSP}(\Gamma')$ are polynomial-time equivalent.

So we may assume that Γ is a core.

Cores and constants

Let $\mathcal{C}_D = \{\{a\} \mid a \in D\}$.

Proposition

Let Γ be a core over D. Then $\textsc{Min CSP}(\Gamma, \mathcal{C}_D)$ is polynomial-time reducible to $\textsc{Min CSP}(\Gamma)$.

The endomorphisms of Γ are encoded as the optimal solutions to an instance I of $\textsc{Min CSP}(\Gamma)$ with

- variables $X_D = \{x_a | a \in D\}$, and
- sum $\sum_{f_i \in \Gamma} \sum_{\mathbf{a} \in f_i^{-1}(0)} f_i(\mathbf{x_a})$, where $\mathbf{x_a} = (x_{a_1}, \ldots x_{a_k}) \in X_D^k$.

When Γ is a core, and $\mathbf{d} = (d_1, \ldots, d_n)$ is an enumeration of D, we have

$$\{(f(d_1), \ldots, f(d_n)) \mid f \in \mathrm{Aut}(\Gamma)\} = \pi_{\mathbf{x_d}} \mathrm{Optsol}(I) \in \langle \Gamma \rangle_w .$$

Add a copy of I to the instance of $\textsc{Min CSP}(\Gamma, \mathcal{C}_D)$, identify variables, solve, and apply an inverse automorphism. □

Outline

Expressive power

We say that a function $h : D^k \to \mathbb{Q}_{\geq 0}$ is **expressible** over (Γ, Δ) if there is an instance I of $\mathrm{VCSP}(\Gamma, \Delta)$ and a sequence of variables $(v_1, \dots, v_k)$ such that

$$h(a_1, \dots, a_k) = \mathrm{Opt}(I \cup \{v_i = a_i\})$$

for all $(a_1, \dots, a_k) \in D^k$, and h is a total function.

Let $\langle \Gamma, \Delta \rangle_{fn}$ denote the set of all such functions.

Proposition

Let $\Gamma' \subseteq \langle \Gamma, \Delta \rangle_{fn}$ be a finite subset. Then $\mathrm{VCSP}(\Gamma', \Delta)$ is polynomial-time reducible to $\mathrm{VCSP}(\Gamma, \Delta)$.

Proposition

If (f, g) is a binary multimorphism of (Γ, Δ), then it is also a multimorphism of $\langle \Gamma, \Delta \rangle_{fn}$.

Graph of partial multimorphisms

We now proceed as follows.

- We first define a graph G which encodes the binary multimorphisms of the binary functions in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ in certain independent sets. (This is a slight extension of a construction by Kolmogorov and Živný (2010) for conservative finite-valued VCSP.)

Graph of partial multimorphisms

We now proceed as follows.

- We first define a graph G which encodes the binary multimorphisms of the binary functions in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ in certain independent sets. (This is a slight extension of a construction by Kolmogorov and Živný (2010) for conservative finite-valued VCSP.)
- From this graph, it is easy to determine whether there is a reduction from $\textsc{Max cut}$.

Graph of partial multimorphisms

We now proceed as follows.

- We first define a graph G which encodes the binary multimorphisms of the binary functions in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ in certain independent sets. (This is a slight extension of a construction by Kolmogorov and Živný (2010) for conservative finite-valued VCSP.)
- From this graph, it is easy to determine whether there is a reduction from $\textsc{Max cut}$.
- Otherwise we can argue, using the graph G, that the binary functions in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ have one of a small number of binary mulitmorphisms which are known to imply tractability.

Graph of partial multimorphisms

We now proceed as follows.

- We first define a graph G which encodes the binary multimorphisms of the binary functions in $\langle\Gamma, \mathcal{C}_D\rangle_{fn}$ in certain independent sets. (This is a slight extension of a construction by Kolmogorov and Živný (2010) for conservative finite-valued VCSP.)
- From this graph, it is easy to determine whether there is a reduction from MAX CUT.
- Otherwise we can argue, using the graph G, that the binary functions in $\langle\Gamma, \mathcal{C}_D\rangle_{fn}$ have one of a small number of binary mulitmorphisms which are known to imply tractability.
- Finally, we need to show that these multimorphisms are in fact multimorphisms of all of Γ.

Graph definition

Definition

Let $G = (V, E)$ be the following undirected graph. V is set of pairs of partial functions $f, g : D^2 \to D$ such that

- f and g are defined on a subset $\{a, b\} \subseteq D$;
- f and g are idempotent and commutative; and
- $\{f(a, b), g(a, b)\} = \{a, b\}$ or $\{f(a, b), g(a, b)\} \cap \{a, b\} = \emptyset$.

$\{(f_1, g_1), (f_2, g_2)\} \in E(G)$ if there is a binary function $h \in \langle \Gamma, \mathcal{C}_D \rangle_{fn}$ s.t.

$$\min\{h(a_1, a_2) + h(b_1, b_2), h(a_1, b_2) + h(b_1, a_2)\} < h(f_1(a_1, b_1), f_2(a_2, b_2)) + h(g_1(a_1, b_1), g_2(a_2, b_2))$$

Note that a and b are allowed to be equal.

Encoding of multimorphisms

For $i = 1, 2$, let $(f_i, g_i) \in V$ be defined on $\{a_i, b_i\}$. Then the edge $\{(f_1, g_1), (f_2, g_2)\} \in E$ means that any pair of operations $(f, g) : D^2 \to D^2$ such that $(f|_{\{a_i,b_i\}}, g|_{\{a_i,b_i\}}) = (f_i, g_i)$ fails to be a multimorphism of $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$.

Lemma

Let $\{(f_{\{a,b\}}, g_{\{a,b\}})\} \subseteq V$ be an independent set in G containing precisely one vertex for each subset $\{a, b\} \subseteq D$. Then, every binary function in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ has the multimorphism (f, g), where f and g are given by

$$f(x, y) = f_{\{x,y\}}(x, y), \text{ and}$$
$$g(x, y) = g_{\{x,y\}}(x, y).$$

NP-hardness

Let $\overrightarrow{xy}$ denote the vertex (f, g) for which $f(x, y) = x$ and $g(x, y) = y$. We say that $\overrightarrow{xy}$ is **conservative**.

Proposition

Let Γ be a core over D. If $\{x, y\} \in \langle \Gamma, \mathcal{C}_D \rangle_w$ and the vertex $\overrightarrow{xy}$ has a self-loop in the graph G, for some $x, y \in D$, then $\textsc{Min}\ \mathrm{CSP}(\Gamma)$ is NP-hard.

Reduce from $\textsc{Max cut}$, via $\textsc{Min}\ \mathrm{CSP}(\Gamma, \mathcal{C}_D)$, to $\textsc{Min}\ \mathrm{CSP}(\Gamma)$. $\square$

If we assume that this proposition does not apply, but $\{x, y\} \in \langle \Gamma, \mathcal{C}_D \rangle_w$, then we can exclude a self-loop on $\overrightarrow{xy}$ in G.

So which 2-element subsets $\{x, y\}$ are in $\langle \Gamma, \mathcal{C}_D \rangle_w$?

Expressible 2-element subsets

Let $e_{ab} : D \to D$ denote the operation $e_{ab}(x) = \begin{cases} b & \text{if } x = a \\ x & \text{otherwise.} \end{cases}$

Lemma

If $e_{ab} \notin \mathrm{End}(\Gamma)$, then $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ contains a unary $\{0,1\}$-valued function u such that $u(a) = 0$ and $u(b) = 1$.

- Let $h : D^k \to \{0,1\} \in \Gamma$ and $a_1, \dots, a_k \in D$ be such that $h(a_1, \dots, a_k) = 0$, but $h(e_{ab}(a_1), \dots, e_{ab}(a_k)) = 1$.
- $a_i = a$ for at least one i; replace these by a variable x.
- $u(x) = h(a_1, \dots, x, \dots, a_k)$ expresses the desired function. $\square$

Thus, when Γ is a core, we have access to a number of unary $\{0,1\}$-valued functions. We can use these functions to determine the 2-element subsets of $\langle \Gamma, \mathcal{C}_D \rangle_w$ when $D = \{a, b, c, d\}$.

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
(1,0,0,0)	(1,0,0,0)	(0,0,0,1)	(0,0,0,1)
(1,0,0,1)	(1,0,0,1)	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	(1,0,0,1)	(1,0,0,1)
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
(1,0,0,0)	(1,0,0,0)	(0,0,0,1)	(0,0,0,1)
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
(1,0,0,0)	(1,0,0,0)	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
(1,0,0,0)	(1,0,0,0)	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

$$(0,0,1,1) + (0,1,0,1) + 2 \cdot (1,0,0,0) = (2,1,1,2)$$
$$(0,0,1,1) + (1,1,0,1) + (1,0,0,0) = (2,1,1,2)$$

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
(1,0,0,0)	(1,0,0,0)	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

$$(0,0,1,1) + (0,1,0,1) + 2 \cdot (1,0,0,0) = (2,1,1,2)$$
$$(0,0,1,1) + (1,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (0,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (1,1,0,1) = (2,1,1,2)$$

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
~~(1,0,0,0)~~	~~(1,0,0,0)~~	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	(1,1,0,1)	(1,0,1,1)	(1,1,0,1)

$$(0,0,1,1) + (0,1,0,1) + 2 \cdot (1,0,0,0) = (2,1,1,2)$$
$$(0,0,1,1) + (1,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (0,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (1,1,0,1) = (2,1,1,2)$$

Expressible 2-element subsets of $\{a, b, c, d\}$

Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
~~(1,0,0,0)~~	~~(1,0,0,0)~~	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
(1,0,1,1)	~~(1,1,0,1)~~	(1,0,1,1)	~~(1,1,0,1)~~

$$(0,0,1,1) + (0,1,0,1) + 2 \cdot (1,0,0,0) = (2,1,1,2)$$
$$(0,0,1,1) + (1,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (0,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (1,1,0,1) = (2,1,1,2)$$
$$2 \cdot (1,0,1,1) + (1,1,0,0) + (0,1,0,1) = (3,2,2,3)$$

Expressible 2-element subsets of $\{a, b, c, d\}$

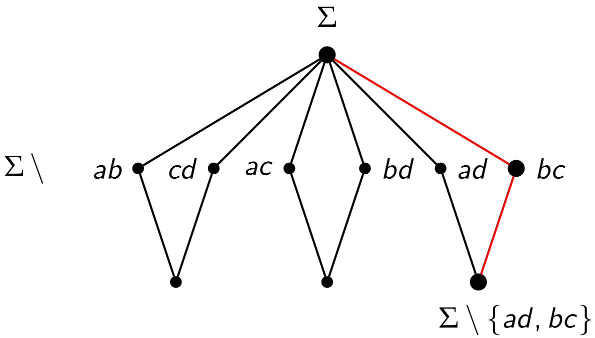
Let Γ be a core and assume that $\{b, c\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$.

e_{ba}	e_{ca}	e_{bd}	e_{cd}
~~(1,0,0,0)~~	~~(1,0,0,0)~~	~~(0,0,0,1)~~	~~(0,0,0,1)~~
~~(1,0,0,1)~~	~~(1,0,0,1)~~	(0,0,1,1)	(0,1,0,1)
(1,0,1,0)	(1,1,0,0)	~~(1,0,0,1)~~	~~(1,0,0,1)~~
~~(1,0,1,1)~~	~~(1,1,0,1)~~	~~(1,0,1,1)~~	~~(1,1,0,1)~~

$$(0,0,1,1) + (0,1,0,1) + 2 \cdot (1,0,0,0) = (2,1,1,2)$$
$$(0,0,1,1) + (1,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (0,1,0,1) + (1,0,0,0) = (2,1,1,2)$$
$$(1,0,1,1) + (1,1,0,1) = (2,1,1,2)$$
$$2 \cdot (1,0,1,1) + (1,1,0,0) + (0,1,0,1) = (3,2,2,3)$$

Expressible 2-element subsets of $\{a, b, c, d\}$

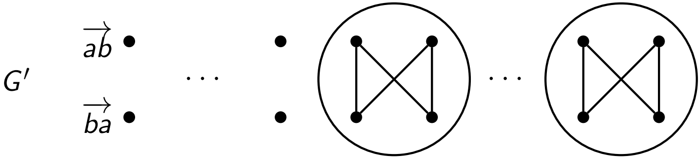
Let Σ be the set of all 2-subsets of $\{a, b, c, d\}$ and let $\Sigma_\Gamma := \Sigma \cap \langle \Gamma, \mathcal{C}_D \rangle_w$.



We will assume that $\Sigma_\Gamma = \Sigma \setminus \{ad, bc\}$, $\Sigma \setminus \{bc\}$, or Σ.

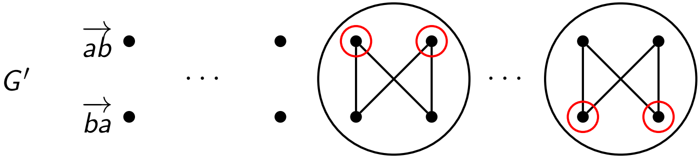
G' bipartite implies submodularity w.r.t. a chain

Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



G' bipartite implies submodularity w.r.t. a chain

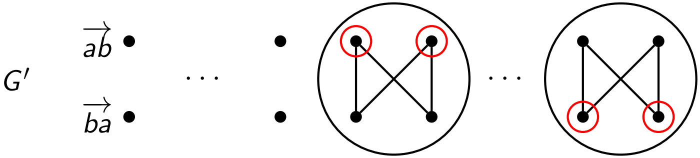
Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



- Let I be the union of one part from each non-trivial component.

G' bipartite implies submodularity w.r.t. a chain

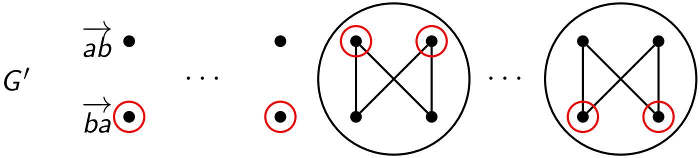
Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



- Let I be the union of one part from each non-trivial component.
- We can show that $\overrightarrow{xy}, \overrightarrow{yz} \in I$ implies $\overrightarrow{xz} \in I$, so I induces a partial order $\prec := \{(x, y) \mid \overrightarrow{xy} \in I\}$.

G' bipartite implies submodularity w.r.t. a chain

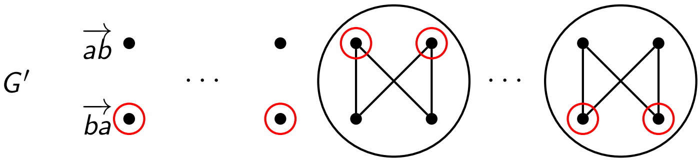
Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



- Let I be the union of one part from each non-trivial component.
- We can show that $\overrightarrow{xy}, \overrightarrow{yz} \in I$ implies $\overrightarrow{xz} \in I$,
  so I induces a partial order $\prec := \{(x, y) \mid \overrightarrow{xy} \in I\}$.
- Extend $\prec$ to a linear order, and add the corresponding vertices to I.

G' bipartite implies submodularity w.r.t. a chain

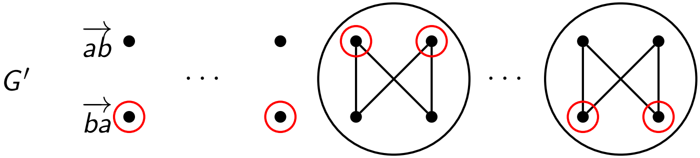
Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



- Let I be the union of one part from each non-trivial component.
- We can show that $\overrightarrow{xy}, \overrightarrow{yz} \in I$ implies $\overrightarrow{xz} \in I$, so I induces a partial order $\prec := \{(x, y) \mid \overrightarrow{xy} \in I\}$.
- Extend $\prec$ to a linear order, and add the corresponding vertices to I.
- Add the vertex on $\{x\}$ to I, for each $x \in D$.

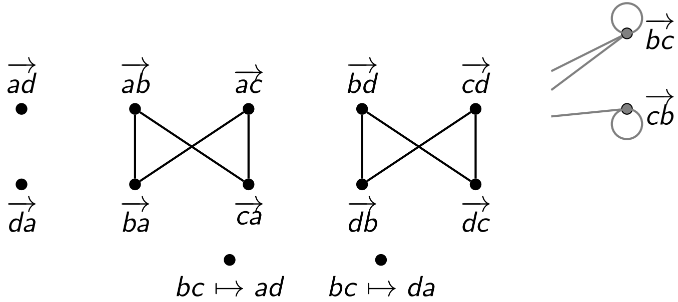
G' bipartite implies submodularity w.r.t. a chain

Let G' be the subgraph of G induced by all conservative vertices $\overrightarrow{xy}$.



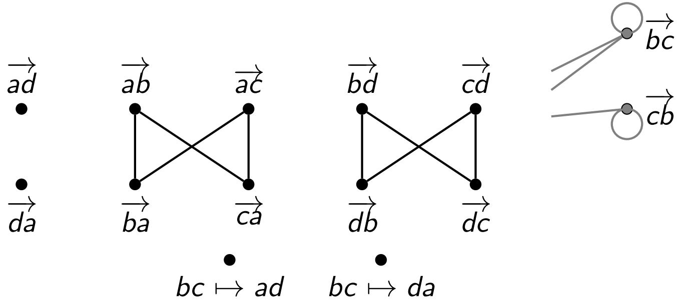
- Let I be the union of one part from each non-trivial component.
- We can show that $\overrightarrow{xy}, \overrightarrow{yz} \in I$ implies $\overrightarrow{xz} \in I$,
  so I induces a partial order $\prec := \{(x, y) \mid \overrightarrow{xy} \in I\}$.
- Extend $\prec$ to a linear order, and add the corresponding vertices to I.
- Add the vertex on $\{x\}$ to I, for each $x \in D$.
- Otherwise $\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



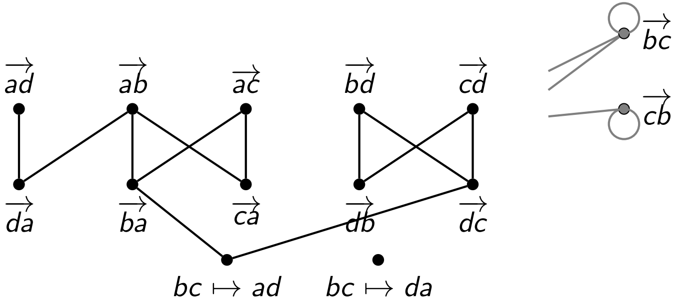
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



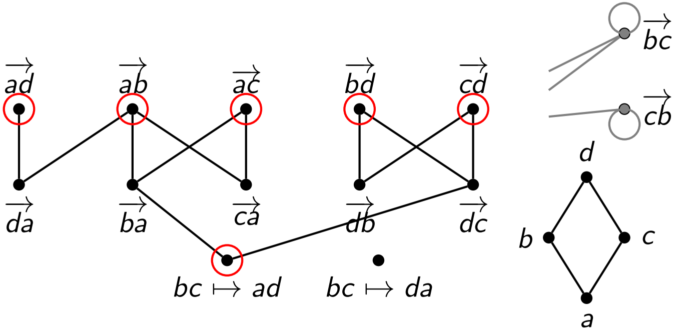
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



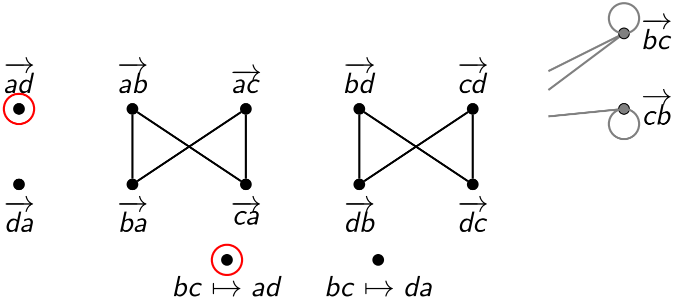
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



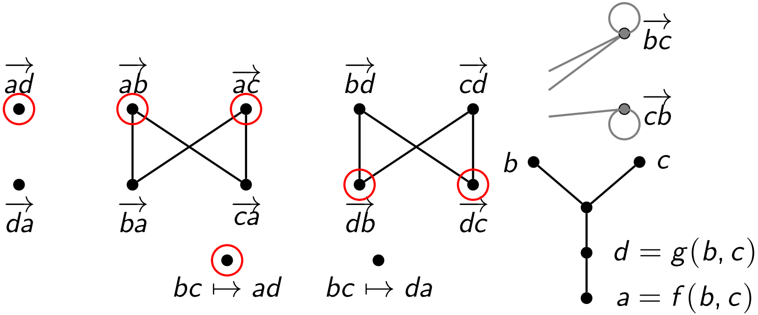
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



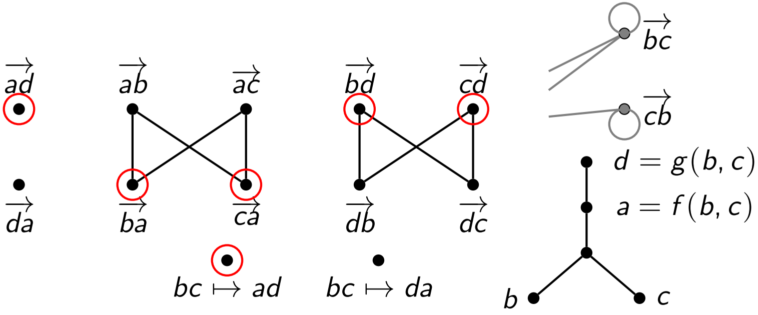
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...
- ... or we can pick $bc \mapsto ad$ and $\overrightarrow{ad}$ and make sure the rest fits.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



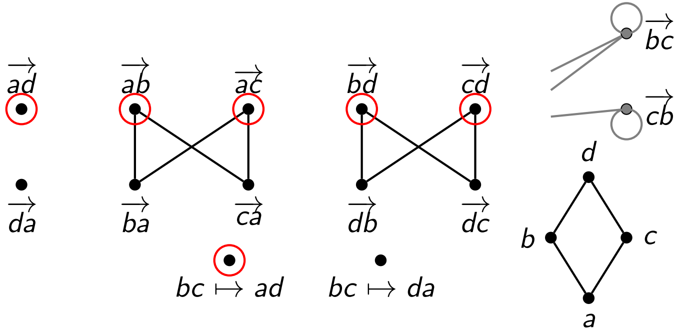
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...
- ... or we can pick $bc \mapsto ad$ and $\overrightarrow{ad}$ and make sure the rest fits.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



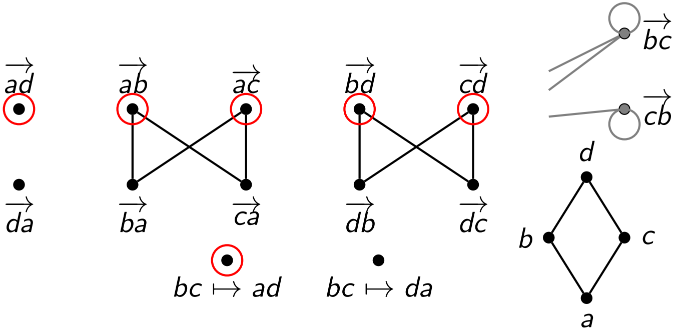
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...
- ... or we can pick $bc \mapsto ad$ and $\overrightarrow{ad}$ and make sure the rest fits.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...
- ... or we can pick $bc \mapsto ad$ and $\overrightarrow{ad}$ and make sure the rest fits.

$\Sigma_\Gamma \subseteq \Sigma \setminus \{bc\}$, 1-defect chains



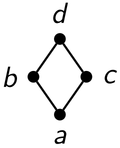
- $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$ is bipartite.
- The vertices of a connected component again induce a partial order.
- Either $bc \mapsto ad$ is connected to a vertex in $G' \setminus \{\overrightarrow{bc}, \overrightarrow{cb}\}$...
- ... or we can pick $bc \mapsto ad$ and $\overrightarrow{ad}$ and make sure the rest fits.
- Add the vertex on $\{x\}$ to I, for each $x \in D$.

Tractable cases, $D = \{a, b, c, d\}$

Assume that for each $x \neq y \in D$, either $\{x, y\} \notin \langle \Gamma, \mathcal{C}_D \rangle_w$ or the vertex $\overrightarrow{xy}$ has no self-loop. Let Γ_{bin} denote the set of at most binary functions in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$.

Up to a permutation of D we have that

- If G' is bipartite, then Γ_{bin} is submodular w.r.t. a chain.
- Otherwise, either Γ_{bin} has a 1-defect chain multimorphism; or
- Γ_{bin} is submodular w.r.t.



Outline

From binary to arbitrary arity

Lemma (Topkis, 1978)

A function $h : D^k \to \mathbb{Q}_{\geq 0}$ is submodular w.r.t. a chain $(D; \wedge, \vee)$ if and only if every binary function obtained from h by replacing any given $k-2$ arguments by constants is submodular on this chain.

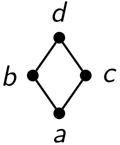
If every binary function in $\langle \Gamma, \mathcal{C}_D \rangle_{fn}$ is submodular w.r.t. a chain, then in particular, every $h \in \Gamma$ fulfils the second part of the lemma, so Γ is submodular w.r.t. this chain.

From binary to arbitrary arity

We can show that the same holds for 1-defect chains.

Lemma

A function $h : D^k \to \mathbb{Q}_{\geq 0}$ has a 1-defect chain multimorphism $\langle f, g \rangle$ if and only if every binary function obtained from h by replacing any given $k-2$ arguments by constants has the multimorphism $\langle f, g \rangle$.

Recall that 

is a 1-defect chain.

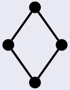
One can also derive the property for this lattice by regarding h as a $2k$-ary function over $\{ac, bd\}$ which is submodular with respect to the chain $ac < bd$.

Classification

To summarise, we have the following.

Theorem

Let Γ be a core with domain $D = \{a, b, c, d\}$.

- *If Γ is submodular w.r.t. a chain on D;*
- *if Γ has a 1-defect chain multimorphism; or*
- *if Γ is submodular w.r.t. a lattice isomorphic to*  ,

then $\textsc{Min CSP}(\Gamma)$ *is tractable.*

Otherwise $\textsc{Min CSP}(\Gamma)$ *is NP-hard.*

Outline

Open problems

Problem

It now seems reasonable to expect that binary idempotent commutative multimorphisms will play a big role in the classification of $\textsc{Min CSP}$.

- *Which of these actually show up?*
  (bisubmodularity does not show up for 3-element $\textsc{Min CSP}$*)*
- *Algorithms!*

Open problems

Problem

It now seems reasonable to expect that binary idempotent commutative multimorphisms will play a big role in the classification of $\textsc{Min CSP}$.

- *Which of these actually show up?* *(bisubmodularity does not show up for 3-element* $\textsc{Min CSP}$*)*
- *Algorithms!*

Problem

Find a general class $\mathcal{C}$ *of binary idempotent commutative multimorphisms such that: a* k*-ary cost function* h *has the multimorphism* $(f, g) \in \mathcal{C}$ *iff every binary function obtained from* h *by replacing any given* $k - 2$ *arguments by constants has the multimorphism* (f, g).