Probabilistic Graphical Models & Applications

Pseudo Boolean Optimization III/III

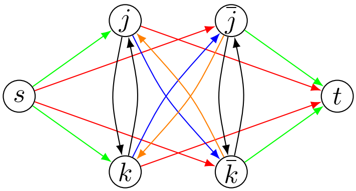
Bjoern Andres and Bernt Schiele

Max Planck Institute for Informatics

Definition 1

For any $n \in \mathbb{N}$ *and any* $c \in C^+_{n2}$, *the* **network** $N = (V, E, s, t, w)$ **of** c *contains the nodes* $V = \{s, t, 1, \bar{1}, \ldots, n, \bar{n}\}$ *and the weighted edges*

for any $c_{\{j\}\emptyset} > 0$	$s\bar{j}, jt$	$w_{s\bar{j}} := w_{jt} := \frac{1}{2}c_{\{j\}\emptyset}$
for any $c_{\emptyset\{j\}} > 0$	$sj, \bar{j}t$	$w_{sj} := w_{\bar{j}t} := \frac{1}{2}c_{\emptyset\{j\}}$
for any $c_{\{j,k\}\emptyset} > 0$	$j\bar{k}, k\bar{j}$	$w_{j\bar{k}} := w_{k\bar{j}} := \frac{1}{2}c_{\{j,k\}\emptyset}$
for any $c_{\{j\}\{k\}} > 0$	$jk, \bar{k}\bar{j}$	$w_{jk} := w_{\bar{k}\bar{j}} := \frac{1}{2}c_{\{j\}\{k\}}$
for any $c_{\emptyset\{j,k\}} > 0$	$\bar{j}k, \bar{k}j$	$w_{\bar{j}k} := w_{\bar{k}j} := \frac{1}{2}c_{\emptyset\{j,k\}}$



Definition 2

For any $n \in \mathbb{N}$, *any* $c \in C^+_{n2}$, *the network* $N = (V, E, s, t, w)$ *of* c *and any* $x \in \{0,1\}^n$, *let* $x' \in \{0,1\}^V$ *such that*

$$x'_s = 1 \tag{1}$$

$$x'_t = 0 \tag{2}$$

$$\forall j \in [n] \quad x'_j = x_j \tag{3}$$

$$\forall j \in [n] \quad x'_{\bar{j}} = 1 - x_j \tag{4}$$

Lemma 1

For any $n \in \mathbb{N}$*, any* $c \in C^+_{n2}$*, the network* $N = (V, E, s, t, w)$ *of* c *and any* $x \in \{0,1\}^n$*, the PBF* $f_c : \{0,1\}^n \to \mathbb{R}$ *defined by* c *is such that*

$$\forall x \in \{0,1\}^n \quad f_c(x) = \sum_{jk \in E} w_{jk} x'_j (1 - x'_k) \ . \tag{5}$$

Proof. Trivial.

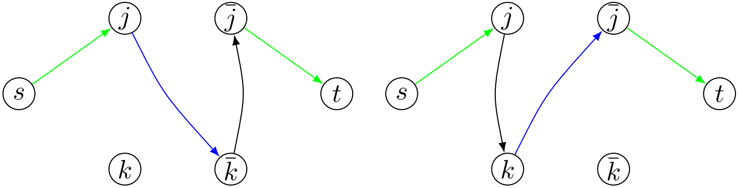
Lemma 2

For any $n \in \mathbb{N}$, *any* $c \in C^{+}_{n2}$, *the network* $N = (V, E, s, t, w)$ *of* c, *any* $m \in \mathbb{N}$ *and any augmenting path* $(sv_1, \ldots, v_m t)$ *in* N, *the path* $(s\bar{v}_m, \ldots, \bar{v}_1 t)$ *is also augmenting.*

Example.

Given path $sj, j\bar{k}, \bar{k}\bar{j}, \bar{j}t$

Conjectured path $sj, jk, k\bar{j}, \bar{j}t$



Posiform corresponding to both paths

$$(1 - x_j) + x_j x_k + (1 - x_k)x_j + (1 - x_j)$$

Lemma 3

For any $n \in \mathbb{N}$, *any* $c \in C^{+}_{n2}$, *the network* $N = (V, E, s, t, w)$ *of* c, *any* st*-flow* $g \in \mathbb{R}^E$ *in* N, *and the posiform* $c' \in C^{+}_{n2}$ *of the residual network of* g,

$$f_c = c_{\emptyset\emptyset} + \varphi_s + f_{c'} \quad . \tag{6}$$

Moreover, the r.h.s. is a posiform which differs from the homogenous posiform c' *only by the additional constant term* $c_{\emptyset\emptyset} + \varphi_s$*.*

Lemma 4

For any $n \in \mathbb{N}$, *any* $f : \{0,1\}^n \to \mathbb{R}$ *and any* $c, c' \in C^+_{n2}(f)$ *such that* $c_{\emptyset\emptyset} < c'_{\emptyset\emptyset}$, *there exists an augmenting path in the network of* c.

Theorem 1

For any $n \in \mathbb{N}$*, any* $f : \{0,1\}^n \to \mathbb{R}$*, any* $c \in C^+_{n2}(f)$ *and the value* $\hat{\varphi}_s \in \mathbb{R}$ *of any maximum* st*-flow in the network of* c*,*

$$r_f = c_{\emptyset\emptyset} + \hat{\varphi}_s \quad . \tag{7}$$

Summary:

- The computation of the floor dual has been reduced to the Maximum st-Flow Problem.

Theorem 2 (strong persistency)

For any $n \in \mathbb{N}$, any $f : \{0,1\}^n \to \mathbb{R}$, any $c \in C^+_{n2}(f)$, any maximum st-flow $g \in \mathbb{R}^E$ in the network of c and the set $S \subseteq V$ of all nodes reachable from s via a path in the residual network of g,

$$\forall \hat{x} \in \operatorname*{argmin}_{x \in \{0,1\}^n} f(x) \quad \forall j \in [n] \quad (j \in S \Rightarrow x_j = 1) \wedge (\bar{j} \in S \Rightarrow x_j = 0) \ .$$

Proof.

- Let $c' \in C^+_{n2}$ be the posiform of the residual network of g.
- Firstly, $c'_{\emptyset\emptyset} = 0$ (by Definition 1)
- Secondly, $\forall j \in [n] : \; j \notin S \vee \bar{j} \notin S$ by the following argument. If $j \in S \wedge \bar{j} \in S$, there exist paths in the residual network

in contradiction to the maximality of the flow g.

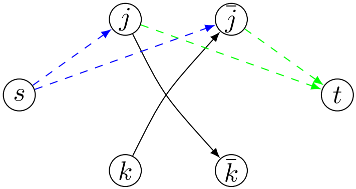
Proof.

- Let $c' \in C^+_{n2}$ be the posiform of the residual network of g.
- Firstly, $c'_{\emptyset\emptyset} = 0$ (by Definition 1)
- Secondly, $\forall j \in [n] : \ j \notin S \vee \bar{j} \notin S$ by the following argument. If $j \in S \wedge \bar{j} \in S$, there exist paths in the residual network

in contradiction to the maximality of the flow g.

Proof.

- Let $c' \in C^{+}_{n2}$ be the posiform of the residual network of g.
- Firstly, $c'_{\emptyset\emptyset} = 0$ (by Definition 1)
- Secondly, $\forall j \in [n] : \; j \notin S \vee \bar{j} \notin S$ by the following argument. If $j \in S \wedge \bar{j} \in S$, there exist paths in the residual network



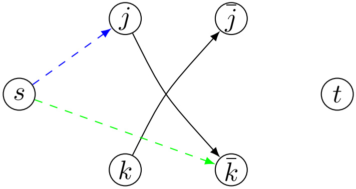
in contradiction to the maximality of the flow g.

- Thirdly,

$$
\begin{aligned}
c'_{\{j,k\}\emptyset} > 0 \ \wedge \ j \in S \quad &\Rightarrow \quad \bar{k} \in S \\
c'_{\{j\}\{k\}} > 0 \ \wedge \ j \in S \quad &\Rightarrow \quad k \in S \\
c'_{\{j\}\{k\}} > 0 \ \wedge \ \bar{k} \in S \quad &\Rightarrow \quad \bar{j} \in S \\
c'_{\emptyset\{j,k\}} > 0 \ \wedge \ \bar{j} \in S \quad &\Rightarrow \quad k \in S
\end{aligned}
$$

(because edges with positive weight cannot direct from a node in S to a node not in S, by definition of S).

Let $c'_{\{j,k\}\emptyset} > 0 \ \wedge \ j \in S$.

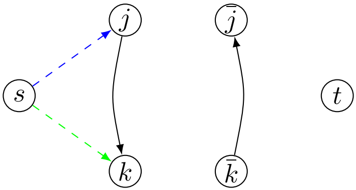


- Thirdly,

$$\begin{aligned} c'_{\{j,k\}\emptyset} > 0 \ \wedge \ j \in S \quad &\Rightarrow \quad \bar{k} \in S \\ c'_{\{j\}\{k\}} > 0 \ \wedge \ j \in S \quad &\Rightarrow \quad k \in S \\ c'_{\{j\}\{k\}} > 0 \ \wedge \ \bar{k} \in S \quad &\Rightarrow \quad \bar{j} \in S \\ c'_{\emptyset\{j,k\}} > 0 \ \wedge \ \bar{j} \in S \quad &\Rightarrow \quad k \in S \end{aligned}$$

(because edges with positive weight cannot direct from a node in S to a node not in S, by definition of S).

Let $c'_{\{j\}\{k\}} > 0 \ \wedge \ j \in S$.

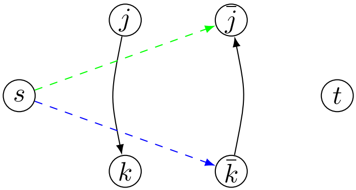


- Thirdly,

$$
\begin{aligned}
c'_{\{j,k\}\emptyset} > 0 \;\wedge\; j \in S &\quad\Rightarrow\quad \bar{k} \in S \\
c'_{\{j\}\{k\}} > 0 \;\wedge\; j \in S &\quad\Rightarrow\quad k \in S \\
c'_{\{j\}\{k\}} > 0 \;\wedge\; \bar{k} \in S &\quad\Rightarrow\quad \bar{j} \in S \\
c'_{\emptyset\{j,k\}} > 0 \;\wedge\; \bar{j} \in S &\quad\Rightarrow\quad k \in S
\end{aligned}
$$

(because edges with positive weight cannot direct from a node in S to a node not in S, by definition of S).

Let $c'_{\{j\}\{k\}} > 0 \;\wedge\; \bar{k} \in S$.

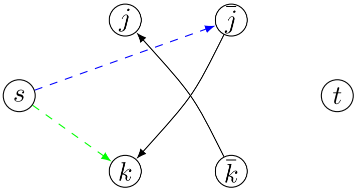


- Thirdly,

$$\begin{aligned}
c'_{\{j,k\}\emptyset} > 0 \;\wedge\; j \in S &\quad\Rightarrow\quad \bar{k} \in S \\
c'_{\{j\}\{k\}} > 0 \;\wedge\; j \in S &\quad\Rightarrow\quad k \in S \\
c'_{\{j\}\{k\}} > 0 \;\wedge\; \bar{k} \in S &\quad\Rightarrow\quad \bar{j} \in S \\
c'_{\emptyset\{j,k\}} > 0 \;\wedge\; \bar{j} \in S &\quad\Rightarrow\quad k \in S
\end{aligned}$$

(because edges with positive weight cannot direct from a node in S to a node not in S, by definition of S).

Let $c'_{\emptyset\{j,k\}} > 0 \;\wedge\; \bar{j} \in S$.



- Therefore, (S', y) with

$$S' = \{j \in [n] \mid j \in S \vee \bar{j} \in S\}$$

and $y : S' \to \{0, 1\}$ such that

$$\forall j \in S' \quad y_j = \begin{cases} 1 & \text{if } j \in S \\ 0 & \text{if } \bar{j} \in S \end{cases}$$

is a **contractor** of c'.

- Fourthly, **weak persistency** holds for (S, y) at the minima of f (because $r_f + f_{c'} = f$).

- Therefore, (S', y) with

$$S' = \{j \in [n] \mid j \in S \vee \bar{j} \in S\}$$

and $y : S' \to \{0, 1\}$ such that

$$\forall j \in S' \quad y_j = \begin{cases} 1 & \text{if } j \in S \\ 0 & \text{if } \bar{j} \in S \end{cases}$$

is a **contractor** of c'.
- Fourthly, **weak persistency** holds for (S, y) at the minima of f (because $r_f + f_{c'} = f$).

- Moreover, $\forall x \in \{0,1\}^n$:

$$f_{c'}(x) = \sum_{jk \in SS} w_{jk} x'_j (1 - x'_k) + \sum_{jk \in SS^C} w_{jk} x'_j (1 - x'_k) + \sum_{jk \in S^C S} w_{jk} x'_j (1 - x'_k) + \sum_{jk \in S^C S^C} w_{jk} x'_j (1 - x'_k)$$

If $\forall j \in S' : x_j = y_j$,

$$f_{c'}(x) = \sum_{jk \in S^C S^C} w_{jk} x'_j (1 - x'_k)$$

Otherwise,

$$f_{c'}(x) > \sum_{jk \in S^C S^C} w_{jk} x'_j (1 - x'_k)$$

Thus, strong persistency holds.

Definition 3

A **lattice** $(S, \preceq)$ *is a set* S, *equipped with a partial order* $\preceq$, *such that any two elements of* S *have an infimum and a supremum w.r.t.* $\preceq$.

Example. $(\{0,1\}^2, \preceq)$ with $\preceq := \{(s,t) \in S \times S \mid s_1 \leq t_1 \wedge s_2 \leq t_2\}$.

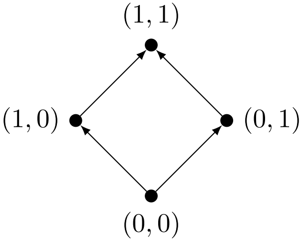
For any $s, t \in \{0,1\}^2$,

$$\sup(s,t) = (\max\{s_1,t_1\}, \max\{s_2,t_2\})$$
$$\inf(s,t) = (\min\{s_1,t_1\}, \min\{s_2,t_2\})$$

Definition 3

A **lattice** $(S, \preceq)$ *is a set* S*, equipped with a partial order* $\preceq$*, such that any two elements of* S *have an infimum and a supremum w.r.t.* $\preceq$*.*

Example. $(\{0,1\}^2, \preceq)$ with $\preceq := \{(s,t) \in S \times S \mid s_1 \leq t_1 \wedge s_2 \leq t_2\}$.



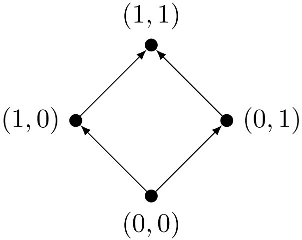
For any $s, t \in \{0,1\}^2$,

$$\sup(s,t) = (\max\{s_1,t_1\}, \max\{s_2,t_2\})$$
$$\inf(s,t) = (\min\{s_1,t_1\}, \min\{s_2,t_2\})$$

Definition 3

A **lattice** $(S, \preceq)$ *is a set* S*, equipped with a partial order* $\preceq$*, such that any two elements of* S *have an infimum and a supremum w.r.t.* $\preceq$*.*

Example. $(\{0,1\}^2, \preceq)$ with $\preceq := \{(s,t) \in S \times S \mid s_1 \leq t_1 \wedge s_2 \leq t_2\}$.



For any $s, t \in \{0,1\}^2$,

$$\sup(s,t) = (\max\{s_1, t_1\}, \max\{s_2, t_2\})$$
$$\inf(s,t) = (\min\{s_1, t_1\}, \min\{s_2, t_2\})$$

Definition 4

A function $f : S \to \mathbb{R}$ *is called* **submodular** *w.r.t. a lattice* $(S, \preceq)$ *iff*

$$\forall s, t \in S \qquad f(\inf(s,t)) + f(\sup(s,t)) \leq f(s) + f(t) \ . \tag{8}$$

Lemma 5

For any $f : \{0,1\}^2 \to \mathbb{R}$, *the following statements are equivalent.*

1. f *is is submodular w.r.t. the the lattice* $(\{0,1\}^2, \preceq)$
2. $f(0,0) + f(1,1) \leq f(1,0) + f(0,1)$
3. *The multi-linear polynomial form*

$$c_\emptyset + c_{\{1\}} x_1 + c_{\{2\}} x_2 + c_{\{1,2\}} x_1 x_2$$

   of f *is such that* $c_{\{1,2\}} \leq 0$.

Proof.

- $f(0,0) + f(1,1) \leq f(1,0) + f(0,1)$ is the only condition in

$$\forall s, t \in S \qquad f(\inf(s,t)) + f(\sup(s,t)) \leq f(s) + f(t)$$

which is not generally true. Thus, (1.) is equivalent to (2.).

- We have

$$\begin{aligned} f(0,0) &= c_\emptyset \\ f(1,0) &= c_\emptyset + c_{\{1\}} \\ f(0,1) &= c_\emptyset \qquad\quad + c_{\{2\}} \\ f(1,1) &= c_\emptyset + c_{\{1\}} + c_{\{2\}} + c_{\{1,2\}} \quad . \end{aligned}$$

Therefore,

$$c_{\{1,2\}} = f(1,1) - f(1,0) - f(0,1) + f(0,0)$$

and thus, (2.) is equivalent to (3.).

Proof.

- $f(0,0) + f(1,1) \leq f(1,0) + f(0,1)$ is the only condition in

$$\forall s,t \in S \qquad f(\inf(s,t)) + f(\sup(s,t)) \leq f(s) + f(t)$$

  which is not generally true. Thus, (1.) is equivalent to (2.).

- We have

$$\begin{aligned} f(0,0) &= c_\emptyset \\ f(1,0) &= c_\emptyset + c_{\{1\}} \\ f(0,1) &= c_\emptyset \qquad\quad + c_{\{2\}} \\ f(1,1) &= c_\emptyset + c_{\{1\}} + c_{\{2\}} + c_{\{1,2\}} \quad . \end{aligned}$$

  Therefore,

$$c_{\{1,2\}} = f(1,1) - f(1,0) - f(0,1) + f(0,0)$$

  and thus, (2.) is equivalent to (3.).

Lemma 6

For every $f : \{0,1\}^2 \to \mathbb{R}$, *there exist unique* $a_0 \in \mathbb{R}$ *and* $a_1, a_{\bar{1}}, a_2, a_{\bar{2}}, a_{12}, a_{\bar{1}2} \in \mathbb{R}_0^+$ *such that*

$$a_1 a_{\bar{1}} = a_2 a_{\bar{2}} = a_{12} a_{\bar{1}2} = 0 \tag{9}$$

and

$$\begin{aligned} \forall x \in \{0,1\}^2 \quad f(x) = \quad & a_0 \\ & + a_1 x_1 + a_{\bar{1}}(1 - x_1) \\ & + a_2 x_2 + a_{\bar{2}}(1 - x_2) \\ & + a_{12} x_1 x_2 + a_{\bar{1}2}(1 - x_1) x_2 \ . \end{aligned} \tag{10}$$

Proof.

- Comparison of (10) with the unique form

$$c_\emptyset + c_{\{1\}}x_1 + c_{\{2\}}x_2 + c_{\{1,2\}}x_1x_2$$

  yields

$$\begin{aligned} a_0 + a_{\bar{1}} + a_{\bar{2}} &= c_\emptyset \\ a_1 - a_{\bar{1}} &= c_{\{1\}} \\ a_2 - a_{\bar{2}} + a_{\bar{1}2} &= c_{\{2\}} \\ a_{12} - a_{\bar{1}2} &= c_{\{1,2\}} \end{aligned} \tag{11}$$

- By these equations (from bottom to top), (9) and c define a uniquely.

Proof.

- Comparison of (10) with the unique form

$$c_\emptyset + c_{\{1\}}x_1 + c_{\{2\}}x_2 + c_{\{1,2\}}x_1x_2$$

  yields

$$\begin{aligned} a_0 + a_{\bar{1}} + a_{\bar{2}} &= c_\emptyset \\ a_1 - a_{\bar{1}} &= c_{\{1\}} \\ a_2 - a_{\bar{2}} + a_{\bar{1}2} &= c_{\{2\}} \\ a_{12} - a_{\bar{1}2} &= c_{\{1,2\}} \end{aligned} \tag{11}$$

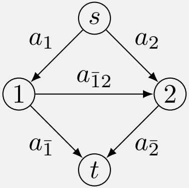
- By these equations (from bottom to top), (9) and c define a uniquely.

Lemma 7 (Kolmogorov and Zabih)

For every **submodular** $f : \{0,1\}^2 \to \mathbb{R}$ *and its unique coefficient* $a_0 \in \mathbb{R}$ *from Lemma 6,*

$$\min_{x \in \{0,1\}^2} f_x - a_0 \tag{12}$$

is equal to the weight of a **minimum** st**-cut** *in the graph below whose edge weights are the (unique, non-negative) coefficients from Lemma 6.*



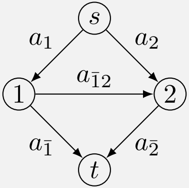
Moreover, f *is minimal at* $\hat{x} \in \{0,1\}^2$ *iff* $\{j \in \{1,2\} | x_j = 0\}$ *is a* **minimum** st**-cutset** *of the above graph.*

Lemma 7 (Kolmogorov and Zabih)

For every **submodular** $f : \{0,1\}^2 \to \mathbb{R}$ *and its unique coefficient* $a_0 \in \mathbb{R}$ *from Lemma 6,*

$$\min_{x \in \{0,1\}^2} f_x - a_0 \tag{12}$$

is equal to the weight of a **minimum** st-**cut** *in the graph below whose edge weights are the (unique, non-negative) coefficients from Lemma 6.*



Moreover, f *is minimal at* $\hat{x} \in \{0,1\}^2$ *iff* $\{j \in \{1,2\} | x_j = 0\}$ *is a* **minimum** st-**cutset** *of the above graph.*

Proof.

- Submodularity of f implies $a_{12} = 0$ in (11), by Lemma 5 and (9).
- Comparison of the four possible minima of f,

$$
\begin{aligned}
f(0,0) &= a_0 + a_{\bar{1}} + a_{\bar{2}} \\
f(1,0) &= a_0 + a_1 + a_{\bar{2}} \\
f(0,1) &= a_0 + a_{\bar{1}} + a_2 + a_{\bar{1}2} \\
f(1,1) &= a_0 + a_1 + a_2 + a_{12} \quad ,
\end{aligned}
$$

with the four possible minimum cuts below proves the Lemma.

Proof.

- Submodularity of f implies $a_{12} = 0$ in (11), by Lemma 5 and (9).
- Comparison of the four possible minima of f,

$$
\begin{aligned}
f(0,0) &= a_0 + a_{\bar{1}} + a_{\bar{2}} \\
f(1,0) &= a_0 + a_1 + a_{\bar{2}} \\
f(0,1) &= a_0 + a_{\bar{1}} + a_2 + a_{\bar{1}2} \\
f(1,1) &= a_0 + a_1 + a_2 + a_{12} \quad ,
\end{aligned}
$$

  with the four possible minimum cuts below proves the Lemma.