

## 12 Umění formulace úloh MIP

Již několikrát jsme viděli, že formulace úloh diskretní optimalizace jako úloh MIP nebývá tak jednoduchá a skrývá v sobě někdy nečekaná úskalí. Jedním z nich je třeba nutnost formulovat exponenciálně mnoho omezení, jiným třeba požadavek formulace "A nebo B", což přímo nejde.

V této lekci si ukážeme hlubší umění správné formulace diskretních úloh v MIP na několika hezkých příkladech.

### Stručný přehled lekce

- Obecnější formulace barevnosti grafu.
- Další triky formulace; jako vyjádření disjunkce či kratší vyjádření TSP.
- Celočíselné polyedry a jejich příklady v kombinatorice.
- Neúplné iterativní formulace úloh.

## 12.1 Více o barevnosti grafu

Pro ilustraci některých pokročilých postupů formulace úloh MIP si zkusme nejprve zapsat problém barevnosti grafu (Příklad 11.7) úplně jinou formulací:

**Příklad 12.1.** *Zformulujme problém barevnosti grafu jako problém MIP, kde barvy přiřazené vrcholům budou přímo (jako čísla) uloženy v reálných proměnných. Zachováme přitom požadavek, aby se hodnoty sousedních barev lišili alespoň o 1.*

**Řešení:** V souladu se zadáním každému vrcholu  $i = 1, 2, \dots, n$  grafu  $G$  přiřadíme reálnou proměnnou  $x_i$ . Nechtě  $\{i, j\}$  je hranou  $G$ . Jak pak v rámci formulace MIP vyjádříme požadavek  $|x_i - x_j| \geq 1$ ? To je, zdá se být, hlavním problémem našeho příkladu.

Trik je následovný – zavedeme dodatečnou binární proměnnou  $z_{i,j}$  pro každou hranu  $\{i, j\}$  a požadavek na barvy rozepíšeme jako

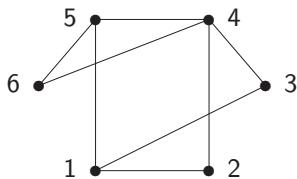
$$\begin{aligned}x_i &\leq x_j - 1 + nz_{i,j}, \\x_i &\geq x_j + 1 - n(1 - z_{i,j}).\end{aligned}$$

(Hodnota proměnné  $z_{i,j}$  nám takto určuje, která z obou barev na hraně  $\{i, j\}$  bude větší.) Skutečně, pokud  $z_{i,j} = 0$ , tak z první nerovnosti plyne  $x_i - x_j \leq -1$ , a pokud  $z_{i,j} = 1$ , tak z druhé nerovnosti plyne  $x_i - x_j \geq 1$ . Opačný směr ekvivalence je také zřejmý.

Nakonec vyjádříme celkový počet použitých barev jako maximum z hodnot nabytých proměnnými  $\vec{x}$  plus 1 (nezapomeňme, že hodnoty  $\vec{x}$  začínají od 0). Takže zapíšeme

$$\begin{aligned} \min y & : \\ \forall i = 1, 2, \dots, n : x_i + 1 & \leq y. \end{aligned}$$

Například pro graf na obrázku zapíšeme úlohu MIP



$$\begin{aligned} \min y & : \\ x_1, x_2, x_3, x_4, x_5, x_6 & \leq y - 1 \\ x_1 & \leq x_2 - 1 + 6z_{12}, & x_1 & \geq x_2 + 1 - 6(1 - z_{12}) \\ x_1 & \leq x_3 - 1 + 6z_{13}, & x_1 & \geq x_3 + 1 - 6(1 - z_{13}) \\ x_2 & \leq x_4 - 1 + 6z_{24}, & x_2 & \geq x_4 + 1 - 6(1 - z_{24}) \\ x_3 & \leq x_5 - 1 + 6z_{35}, & x_3 & \geq x_5 + 1 - 6(1 - z_{35}) \\ x_4 & \leq x_5 - 1 + 6z_{45}, & x_4 & \geq x_5 + 1 - 6(1 - z_{45}) \end{aligned}$$

$$\begin{aligned}
x_1 &\leq x_5 - 1 + 6z_{15}, & x_1 &\geq x_5 + 1 - 6(1 - z_{15}) \\
x_6 &\leq x_5 - 1 + 6z_{65}, & x_6 &\geq x_5 + 1 - 6(1 - z_{65}) \\
x_4 &\leq x_6 - 1 + 6z_{46}, & x_4 &\geq x_6 + 1 - 6(1 - z_{46}) \\
x_1, x_2, x_3, x_4, x_5, x_6 && &\geq 0 \\
z_1, z_2, z_3, z_4, z_5, z_6 && &\in \{0, 1\}.
\end{aligned}$$

Poněkud stranou zatím zůstala důležitá otázka, zda jsme naší úlohou MIP zapsali skutečně problém klasické barevnosti grafu. Vždyť jsme výrazně rozšířili množinu přípustných řešení na reálná čísla. Přesto lze dokázat, že optimálním výsledkem naší úlohy je skutečně klasická (celočíslná) barevnost.  $\square$

Důležitým přínosem nových lineárních formulací kombinatorických problémů je také možnost jejich zobecňování (či relaxace) umožněné novým pohledem. Například výše uvedená formulace grafové barevnosti s reálnými “barvami” přímo motivuje následující (již dříve známou) neceločíselnou relaxaci barevnosti grafu.

**Definice:** *Cirkulární barevnost* grafu  $G$  rozumíme nejmenší reálné číslo  $f > 0$  takové, že vrcholy  $G$  lze zobrazit na kruhový interval délky (tj. obvodu)  $f$  tak, že sousední vrcholy  $G$  mají na obvodu kruhu vzdálenost alespoň 1.

**Příklad 12.2.** *Zformulujme problém cirkulární barevnosti grafu jako problém MIP, kde barvy přiřazené vrcholům budou přímo uloženy v reálných proměnných.*

Na začátek si blíže osvětlíme význam kruhového intervalu. Dvě čísla (body)  $s, t$  jsou na kružnici o obvodu  $f$  ve vzdálenosti alespoň  $d$  pokud zároveň  $|s - t| \geq d$  a  $|s - t| \leq f - d$  (vzdálenosti měřené v obou směrech na kružnici). Podmínka vzdálenosti hodnot sousedních vrcholů  $G$   $|x_i - x_j| \geq 1$  byla zformulována již v Příkladě 12.1. Pro druhou podmínku vzdálenosti na kruhu se v tomto případě ukazuje jako lepší ji přímo neformulovat, ale místo toho implicitně definovat obvod  $f$  kruhového intervalu jako minimum z hodnot  $|x_i - x_j| + 1$  přes všechny hrany  $G$ .

**Řešení:** S výhodou využijeme již zapsané podmínky z Příkladu 12.1

$$\begin{aligned}\forall \{i, j\} \in E(G) : x_i &\leq x_j - 1 + nz_{i,j} \\ x_i &\geq x_j + 1 - n(1 - z_{i,j}) \\ z_{i,j} &\in \{0, 1\}\end{aligned}$$

a přidáme účelovou funkci spolu s podmínkami

$$\begin{aligned}\min y & : \\ \forall i = 1, 2, \dots, n : x_i - x_j + 1 &\leq y \\ x_j - x_i + 1 &\leq y.\end{aligned} \tag{1}$$

Pokud cirkulární barevnost  $G$  je  $f$ , tak výše zapsaný systém podmínek má přípustné řešení s  $y = f$  dané tímto cirkulárním obarvením.

Naopak pokud vezmeme přípustné řešení výše zapsaného systému podmínek a položíme  $f = y$ , pak snadno přiřadíme vrcholu  $i$  grafu  $G$  barvu  $c_i = x_i \bmod (f + 1)$ . Potom bude splněno  $|c_i - c_j| \geq 1$  a také  $|c_i - c_j| \leq y - 1 = f - 1$  pro každou hranu  $\{i, j\} \in E(G)$ , neboli se bude jednat o platné cirkulární obarvení.  $\square$

## 12.2 Užitečné triky formulace

### Modelování disjunkce

Trik použité v Příkladě 12.1 lze zobecnit na modelování jakéhokoliv počtu *disjunktivních podmínek* v úlohách MIP. Stručně řečeno, dokážeme vyjádřit podmínky ve stylu “platí toto nebo toto nebo toto...” (nebo zde není exkluzivní).

**Tvrzení 12.3.** *Mějme úlohu MIP s omezenou množinou přípustných řešení  $Q$ . Necht'  $P_1, \dots, P_k$  jsou polyedry určené nerovnostmi  $C^j \cdot \vec{x} \leq \vec{d}^j$  pro  $j = 1, \dots, k$  (kde vektor  $\vec{x}$  zahrnuje všechny proměnné úlohy). Pak úlohu s množinou přípustných řešení  $Q \cap (P_1 \cup \dots \cup P_k)$  lze vyjádřit jako úlohu MIP.*

**Důkaz:** Zvolme nejprve dostatečně velkou konstantu  $K$ . Zavedeme nové binární proměnné  $z_1, \dots, z_k$  spolu s přidáními vztahy

$$\begin{aligned} z_1 + z_2 + \dots + z_k &= k - 1 \\ \forall j = 1, 2, \dots, k: C^j \cdot \vec{x} &\leq \vec{d}^j + K \cdot \vec{1} \cdot z_j \end{aligned} \quad (2)$$

Pokud vektory  $\vec{x} \in Q$  a  $\vec{z}$  splňují vztahy (2), tak jedno  $z_i = 0$ , neboli  $\vec{x}$  splňuje  $C^i \cdot \vec{x} \leq \vec{d}^i$  a  $\vec{x} \in Q \cap P_i$ .

Na druhou stranu pokud  $\vec{x} \in Q \cap (P_1 \cup \dots \cup P_k)$ , tak pro některé  $i$  je  $\vec{x} \in Q \cap P_i$ . Zvolíme  $z_i = 0$  a  $z_j = 1$  pro  $j \neq i$ . Pak z předpokladu  $\vec{x} \in P_i$  bude platit  $C^i \cdot \vec{x} \leq \vec{d}^i$  a navíc pro vhodnou volbu konstanty  $K$  bude  $C^j \cdot \vec{x} \leq \vec{d}^j + K \cdot \vec{1}$  platit pro všechna  $j = 1, \dots, k$ . Tudíž (2) bude splněno.  $\square$

## Krátká formulace TSP

Další moc hezký trik formulace úloh MIP je ukázán na zápise problému TSP polynomiálně mnoha nerovnostmi. Přirozená formulace v Příkladu 11.9 vyžaduje exponenciální počet nerovností a pozdější objevení výrazně kratšího zápisu úlohy bylo docela velkým překvapením.

**Tvrzení 12.4.** *Problém obchodního cestujícího na orientovaném grafu  $G$  s  $n$  vrcholy a ohodnocením hran  $w : E(G) \rightarrow \mathbf{R}^+$  lze vyjádřit následujícími **polynomiálně mnoha** nerovnostmi jako úlohu MIP s binárními proměnnými  $z_{i,j}$  a reálnými  $u_i$ :*

$$\begin{aligned} \forall k = 1, 2, \dots, n : \quad \sum_{i:(i,k) \in E(G)} z_{i,k} &= 1 \\ \forall k = 1, 2, \dots, n : \quad \sum_{i:(k,i) \in E(G)} z_{k,i} &= 1 \\ \forall (i,j) \in E(G), i, j > 1 : \quad u_i - u_j + n \cdot z_{i,j} &\leq n - 1 \\ &\vec{u} \geq 0 \\ &\vec{z} \in \{0, 1\}^* \end{aligned} \quad (3)$$

**Důkaz:** Předpokládejme pro spor, že přípustné hodnoty proměnných  $z_{i,j} = 1$  nevyjadřují v grafu  $G$  Hamiltonovský cyklus. Jelikož do každého vrcholu přichází jedna hrana se  $z_{i,j} = 1$  a jedna taková odchází, množina hran se  $z_{i,j} = 1$  je tvořena více cykly a alespoň jeden z nich, označme jej  $C$ , neprochází vrcholem 1. Sečteme-li nerovnosti (3) přes hrany cyklu  $C$ , proměnné  $u_i$  se vyruší a vyjde

$$|V(C)| \cdot n \cdot 1 \leq (n - 1) \cdot |V(C)| ,$$



což je spor.

Naopak máme dokázat, že každý Hamiltonovský cyklus  $C$  v  $G$  (tj. procházející přes všechny vrcholy) je přípustný. Necht' vrcholy  $G$  následují v cyklu  $C$  v pořadí  $v_1 = 1, v_2, v_3, \dots, v_n$ . Položíme  $u_i = l$ , kde  $i = v_l$  (tj. podle pořadí na cyklu). Pak pro každou hranu  $(i, j) \in E(C)$ ,  $i, j > 1$  platí  $u_i - u_j = l - (l + 1) = -1$ , takže

$$u_i - u_j + n \cdot 1 \leq -1 + n \cdot 1 = n - 1.$$

Pro kteroukoliv jinou hranu přirozeně platí  $u_i - u_j + n \cdot 0 \leq n - 1$ . Tím jsme dokončili důkaz správnosti formulace (3) pro problém TSP.  $\square$

**Poznámka:** Naneštěstí se uvedená krátká formulace ukazuje jako **velmi nevhodná** k praktickému řešení problému TSP, neboť její lineární relaxace jen velmi hrubě aproximuje množinu přípustných celočíselných řešení. Proto má tento výsledek povětšinou jen teoretický význam.

## 12.3 Celočíselné polyedry

Pro úvod začneme s přirozenou (a trochu upravenou) formulací problému toků v síti jako úlohy LP.

**Definice:**  $A$  je *vrcholově–hranová incidenční matice* orientovaného grafu  $G$ , pokud její složky (indexované vrcholy  $\times$  hranami) jsou

$$a_{i,(i,j)} = 1 \text{ pro } (i,j) \in E(G), \quad a_{i,(j,i)} = -1 \text{ pro } (j,i) \in E(G), \quad a_{i,e} = 0 \text{ jinak.}$$

Mějme síť  $(G, z, s, \vec{c})$ , kde  $G$  je orientovaný graf,  $z$  zdroj,  $s$  stok a  $\vec{c}$  udává kapacity jednotlivých hran. (Definice 2.1) Doplňme graf  $G$  o "zpětnou" hranu ze stoku  $s$  do zdroje  $z$  bez omezení kapacity a přiřadme proměnné  $x_{(i,j)}$  tokům na jednotlivých hranách. Pak úlohu hledání maximálního toku zapíšeme jako:

$$\begin{aligned} \max x_{(s,z)} & : \\ \begin{pmatrix} \mathbf{A} \\ \mathbf{I} \end{pmatrix} \cdot \vec{x} & = \begin{pmatrix} 0 \\ \vec{c} \end{pmatrix} \\ \vec{x} & \geq 0 \end{aligned}$$

**Definice:** Polyedr  $P$  je *celočíselný*, pokud každá jeho neprázdná stěna obsahuje vektor se všemi celočíselnými složkami.

Omezený polyedr  $P$  je celočíselný podle této definice, pokud každý jeho vrchol je vektorem s celočíselnými složkami.

**Fakt:** Z celočíselnosti řešení Algoritmu 2.6 pro tok v síti v případě celočíselných kapacit  $\vec{c}$  (nepřímo) vyplývá, že polyedr přípustných toků v dané síti je celočíselný.

Uvedený fakt není náhodný, je ve skutečnosti jen speciálním případem širšího fenoménu popsaného následně.

## Totálně unimodulární matice

**Definice:** Řekneme, že matice  $A$  je *totálně unimodulární* (TU) pokud každá její čtvercová podmatice má determinant 0, 1 nebo  $-1$ .

Není těžké zjistit, že složky totálně unimodulární matice mohou být jen 0, 1 nebo  $-1$ , ale to není dostačující. Například jednotková matice  $I$  je totálně unimodulární a taková je třeba i matice

$$\begin{pmatrix} -1 & 1 & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 1 & 0 & 0 & 1 & -1 \end{pmatrix},$$

kteřá má mezi TU maticemi výsadní postavení, jehož význam přesahuje rámec našeho textu.

Význam totálně unimodulárních matic v oblasti celočíselných mnohostěňů je naznačen následující vlastností, která přímo plyne ze známého “determinantového” vztahu pro inverzní matici.

**Fakt:** Inverze regulární totálně unimodulární matice je celočíselná matice.

Z uvedeného faktu a znalosti souvislosti bázeckých řešení LP s vrcholy polyedru (Lema 6.3) není těžké odvodit, že pro TU matici  $A$  a celočíselné  $\vec{b}$  je polyedr určený vztahy  $A \cdot \vec{x} \leq \vec{b}$ ,  $\vec{x} \geq 0$  celočíselný. V obecnosti dokonce lze tvrdit (zde bez důkazu):

**Věta 12.5.** *Matice  $A$  je totálně unimodulární, právě když pro každé  $\vec{b} \in \mathbb{Z}^m$  je polyedr určený vztahy  $A \cdot \vec{x} \leq \vec{b}$ ,  $\vec{x} \geq 0$  celočíselný.*

Souvislost totálně unimodulárních matic s toky v sítích se ukáže následovně.

**Tvrzení 12.6.** *Nechť  $A$  je vrcholově–hranová incidenční matice orientovaného grafu  $G$ . Pak  $A$  je totálně unimodulární.*

**Důkaz:** Vezměme čtvercovou podmatici  $C$  v  $A$ . Tvrzení je jasné, pokud  $C$  má jediný sloupec. Dále postupujeme indukcí podle velikosti  $C$ .

Pokud  $C$  je singulární, má determinant 0. Tato možnost mimo jiné nastává, pokud každý sloupec  $C$  obsahuje obě nenulové souřadnice (1 a  $-1$ ) z příslušného sloupce matice  $A$ . Jinak najdeme sloupec v  $C$  mající právě jednu nenulovou souřadnici. Podle tohoto sloupce determinant rozvineme na jediný poddeterminant, jehož hodnota je  $\pm 1$  z indukčního předpokladu, tudíž to samé platí pro  $|C|$ .  $\square$

## Párování v grafu

Mimo TU matice existují i další zajímavé případy celočíselných polyedrů vycházejících z grafových problémů. Jedním z nich je tzv. mnohostěn párování. Připomínáme, že *párování* v grafu je taková podmnožina hran, ve které žádné dvě hrany nesdílejí vrchol. Párování je *perfektní*, pokud obsahuje všechny vrcholy grafu. (Mimo jiné musí mít takový graf sudý počet vrcholů.)

**Definice:** Nechtě  $\vec{x} \in \{0, 1\}^*$  je charakteristický vektor podmnožin hran grafu  $G$ . *Mnohostěn párování* grafu  $G$  je konvexní obal všech takových charakteristických vektorů  $\vec{x}$ , které odpovídají párováním v  $G$ . *Mnohostěn perfektních párování* grafu  $G$  je konvexní obal charakteristických vektorů perfektních párování v  $G$ .

Vzpomeňme si na Příklad 11.4 ukazující IP formulaci problému párování v grafu. Je zřejmé, že pokud bychom k oné formulaci udělali LP-relaxaci, dostaneme mnohem více řešení než jen ta z mnohostěnu párování.

Například ohodnocení všech proměnných hran na liché kružnici číslem  $\frac{1}{2}$  vyhovuje nerovnostem Příkladu 11.4, ale takové ohodnocení zřejmě nemá nic společného se skutečnými párováními. Jak vidíme při bližším pohledu, problémy dělají liché podmnožiny vrcholů indukující podgrafy s plně "nasycenými" ohodnocenými hranami, což je věc nemožná pro skutečné párování grafu.

**Věta 12.7.** *Mnohostěn perfektních párování grafu  $G$  s  $n$  vrcholy je vyjádřen následujícími vztahy:*

$$\forall i = 1, 2, \dots, n : \quad \sum_{e \in E(G) : e = \{i, j\}} x_e = 1 \quad (4)$$

$$\forall W \subseteq V(G), |W| \text{ liché} : \quad \sum_{e \in E(G) : e = \{i, j\} \subseteq W} x_e \leq \lfloor |W|/2 \rfloor \quad (5)$$

$$\vec{x} \geq 0$$

*Mnohostěn párování grafu  $G$  je vyjádřen stejnými vztahy, jen vztah (4) je zapsán s nerovností  $\leq$  místo  $=$ .*

**Důkaz** (náznak): Nejprve si krátce ukažme, jak z platnosti věty pro perfektní párování plyne její platnost pro všechna párování:

.....

Nyní dokážeme, že vztahy (4),(5) plně popisují mnohostěn perfektních párování grafu  $G$ . V jednom směru potřebujeme dokázat, že každé perf. párování splňuje vztahy (4) a (5). Pro (4) to platí z definice perf. párování. Pokud  $W$  je lichá podmnožina vrcholů, tak aspoň jeden z vrcholů musí zůstat uvnitř  $W$  nespárovaný, a proto plyne (5).

Naopak vezměme vektor  $\vec{x}$ , který splňuje výše uvedené vztahy a zároveň je krajním bodem mnohostěnu perfektních párování. Předpokládejme pro spor, že  $\vec{x}$  není celočíselný. Hranám  $e$  grafu  $G$ , které mají necelé hodnoty  $x_e$ , řekněme *šedé hrany*. Pro šedou kružnici  $C$  sudé délky v  $G$  nazvěme  $\varepsilon$ -alternací  $\vec{x}$  takové ohodnocení  $\vec{x}'$ , které na  $i$ -té hraně v pořadí na kružnici  $C$  nabývá  $x'_{e_i} = x_{e_i} + (-1)^i \varepsilon$  a mimo  $C$  zůstává stejné jako  $\vec{x}$ .

**Tvrzení 12.8.** *Za výše uvedených předpokladů v  $G$  existuje sudá šedá kružnice  $C$  a konstanta  $\delta > 0$  takové, že všechny  $\varepsilon$ -alternace  $C$  jsou přípustné ve vztazích (4),(5) pro  $-\delta < \varepsilon < \delta$ .*

Tvrzení dokazujeme indukcí podle  $n$ , použitím kontrakcí lichých “nasyčených” podmnožin. . .

Nyní již snadno dokončíme důkaz celé věty. Vektor  $\vec{x}$  je totiž konvexní kombinací svých  $\varepsilon$ -alternací a  $-\varepsilon$ -alternací, které jsou přípustné v úloze pro dostatečně malé  $\varepsilon > 0$ . To je ale ve sporu s předpokladem, že  $\vec{x}$  je krajním bodem polyedru.  $\square$

## Mnohostěn perfektního grafu

Jiný známý příklad kombinatorických celočíselných mnohostěňů vychází z popisu klik v jistých grafech. Připomínáme, že *klika* v grafu je jiný název pro inkluzí maximální úplný podgraf.

**Definice:** Graf  $G$  je *perfektní*, pokud v každém jeho indukovaném podgrafu je barevnost rovna velikosti největší kliky.

Nepletme si pojem perfektního grafu s perfektním párováním, není mezi nimi žádná souvislost. Příklady perfektních grafů jsou bipartitní grafy (barevnost 2), intervalové grafy (určené průniky intervalů na přímce), nebo tzv. chordální grafy (bez indukovaných kružnic delších než 3).

**Definice:** Necht  $K$  je incidenční matice klik a vrcholů perfektního grafu  $G$ . (Tj.  $K_{i,j} = 1$  právě když vrchol  $j$  náleží do  $i$ -té kliky.) Pak *mnohostěn perfektního grafu*  $G$  je určený vztahy

$$\begin{aligned} K \cdot \vec{x} &\leq 1 \\ \vec{x} &\geq 0. \end{aligned}$$

**Věta 12.9.** *Mnohostěn perfektního grafu je vždy celočíselný.*



## 12.4 Neúplné formulace úloh MIP

Iterativní vylepšování formulace MIP novými podmínkami na základě předchozích relaxovaných řešení.....