1 Relaksacja Lagrange'a

- Podejście oparte na dekompozycji.
  - Łatwe ograniczenia.
  - "Niewygodne ograniczenia".

  "Niewygodne ograniczenia" dodaje się do funkcji celu.
- Otrzymanie dolnego ograniczenia na optymalną wartość funkcji celu problemu minimalizacyjnego.
- W wielu przypadkach dolne ograniczenie jest bliskie wartości rozwiązania optymalnego.

Problem najkrótszej ścieżki z ograniczeniem czasowym (Constrained Shortest Path) CSP

Dana jest sieć $G = (V, A)$ oraz dwa wyróżnione wierzchołki s i t. Z każdym łukiem sieci $(i, j) \in A$ związane są dwa atrybuty **koszt** c_{ij} i **czas przejazdu** t_{ij}.
Chcemy znaleźć najtańszą ścieżkę od s do t, której czas przejazdu nie przekracza zadanego T.

Model CSP

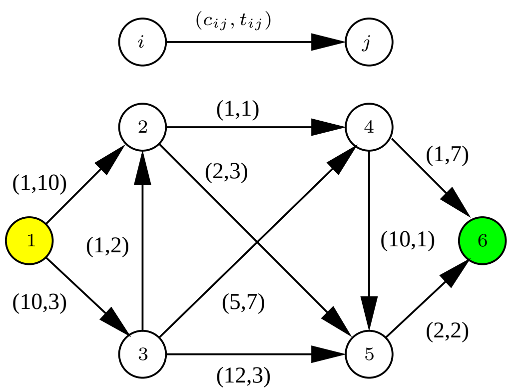
$$z^* = \min \sum_{(i,j)\in A} c_{ij}x_{ij}$$

$$\sum_{j\in S(i)} x_{ij} - \sum_{j\in P(i)} x_{ji} = \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$

$$\sum_{\mathbf{(i,j)\in A}} \mathbf{t_{ij}x_{ij}} \leqslant \mathbf{T} \quad \text{Niewygodne ograniczenie,}$$

$$x_{ij} \in \{0, 1\} \quad (i, j) \in A.$$

Należy znaleźć najtańszą ścieżkę od wierzchołka 1 do 6, której czas przejazdu nie przekracza 10.



- Problem najkrótszej ścieżki jest wielomianowo rozwiązywalny (łatwy).
- Problem najkrótszej ścieżki z ograniczeniem czasowym jest NP-trudny.

Mówimy, że problem optymalizacyjny P' jest *relaksacją* problemu optymalizacyjnego P, jeżeli P' powstaje z P przez eliminacje jednego lub więcej ograniczeń.

Pierwsze podejście do CSP – odrzucenie ograniczenia

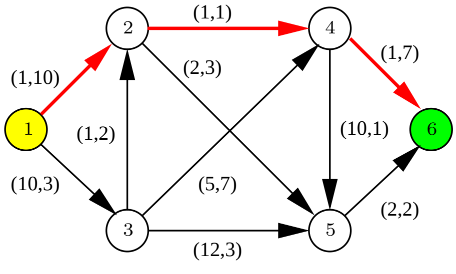
Rozwiązać problem **CSP** bez "niewygodnego ograniczenia".

$$\min \sum_{(i,j)\in A} c_{ij}x_{ij}$$

$$\sum_{j\in S(i)} x_{ij} - \sum_{j\in P(i)} x_{ji} = \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$

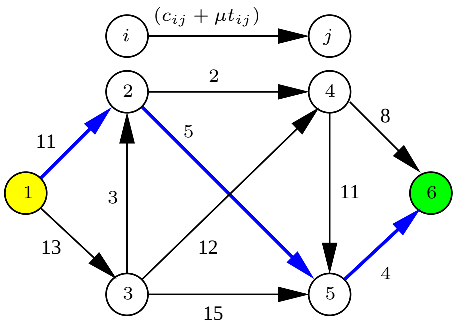
$$x_{ij} \in \{0,1\} \quad (i,j) \in A.$$

Jeżeli rozwiązanie spełnia ograniczenie czasowe, wówczas jest on rozwiązaniem optymalnym problemu **CSP**.



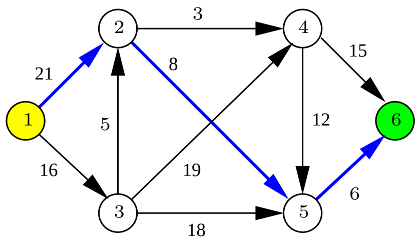
Otrzymana najkrótsza ścieżka ma długość **3** i czas przejazdu równy **18** – nie jest rozwiązanie dopuszczalnym **CSP**. Jej długość jest dolnym ograniczeniem na długość optymalnej ścieżki dla **CSP**.

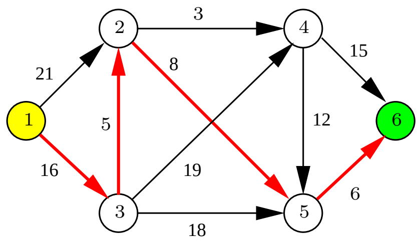
Drugie podejście do CSP – wprowadzenie kar



$\mu = 1$. Koszt optymalnej ścieżki dla zmodyfikowanych wag jest **5**, a czas przejazdu jest równy **15**. Zatem ścieżka jest niedopuszczalna.

Po zwiększeniu μ, $\mu = 2$ otrzymujemy dwie ścieżki, których zmodyfikowany koszt, $\sum_{(i,j)\in p}(c_{ij} + \mu t_{ij})$, jest **35**.





1. $1-2-5-6$ ma czas przejazdu 15 (niedopuszczalna dla **CSP**).
2. $1-3-2-5-6$ ma czas przejazdu 10 jest dopuszczalna dla **CSP** oraz jest rozwiązaniem optymalnym dla **CSP**.

Analiza parametryczna

μ	$\sum_{(i,j)\in p}(c_{ij} + \mu t_{ij})$	$\sum_{(i,j)\in p} c_{ij}$	$\sum_{(i,j)\in p} t_{ij}$	$\sum_{(i,j)\in p}(c_{ij} + \mu t_{ij}) - \mu 10$
0	3	3	18	3
1	20	5	15	10
2	35	5	15	15
3	45	15	10	15
4	55	15	10	15
5	64	24	8	14
6	72	24	8	12
7	80	24	8	10
8	88	24	8	8

Oznaczmy przez $c_p = \sum_{(i,j)\in p} c_{ij}$ koszt ścieżki p oraz przez $t_p = \sum_{(i,j)\in p} t_{ij}$ czas przejazdu przez ścieżkę p. Dla ustalonego μ i ścieżki p zmodyfikowany koszt jest równy

$$c_p(\mu) = \sum_{(i,j)\in p} (c_{ij} + \mu t_{ij}) = c_p + \mu t_p.$$

Własność 1 (Zasada ograniczania (bounding principle)). *Niech $\mu \geqslant 0$ i niech p^* będzie najtańszą ścieżką w sieci ze zmodyfikowanymi kosztami $c_{ij} + \mu t_{ij}$, wówczas $c_{p^*}(\mu) - \mu T$ jest dolnym ograniczeniem długość ścieżki* **CSP**.

Dowód. Ścieżka p^* jest najtańsza, zatem

$$c_{p^*}(\mu) \leqslant c_p + \mu t_p \quad \text{dla dowolnej ścieżki } p$$

w szczególności ścieżki **CSP** p' $(t_{p'} \leqslant T)$.

$$c_{p^*}(\mu) \leqslant c_{p'} + \mu t_{p'} \leqslant c_{p'} + \mu T.$$

□

W kierunku relaksacji Lagrange'a

$$z^* = \min \sum_{(i,j)\in A} c_{ij}x_{ij}$$

$$\sum_{j\in S(i)} x_{ij} - \sum_{j\in P(i)} x_{ji} = \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$
$$\sum_{(i,j)\in A} t_{ij}x_{ij} \leqslant T$$
$$x_{ij} \in \{0,1\} \quad (i,j) \in A.$$

Dla każdego $\mu \geqslant 0$ zastępujemy funkcję celu **CSP** przez

$$\begin{aligned} z(\mu) &= \min \sum_{(i,j)\in A} c_{ij}x_{ij} + \mu\Big(\sum_{\mathbf{(i,j)\in A}} \mathbf{t_{ij}x_{ij} - T}\Big) \\ &= \sum_{(i,j)\in A} (c_{ij} + \mu t_{ij})x_{ij} - \mu T. \end{aligned}$$

Zastosowanie relaksacji Lagrange'a do CSP

$$L(\mu) = \min \sum_{(i,j)\in A} (c_{ij} + \mu t_{ij})x_{ij} - \mu T$$

$$\sum_{j\in S(i)} x_{ij} - \sum_{j\in P(i)} x_{ji} = \begin{cases} 1 & i = s, \\ -1 & i = t, \\ 0 & \text{w przeciwnym przypadku,} \end{cases}$$
$$x_{ij} \in \{0,1\} \quad (i,j) \in A.$$

Twierdzenie 1. *Niech p^* będzie najtańszą ścieżką w sieci ze zmodyfikowanymi kosztami $c_{ij} + \mu t_{ij}$, dla której*
$c_{p^*}(\mu) = c_{p^*} + \mu T$, to p^* jest optymalna.

Dowód. $c_{p^*}(\mu) - \mu T = L(\mu) \leqslant z^* \leqslant c_{p^*} = c_{p^*}(\mu) - \mu T.$ □

Ograniczenia równościowe

$$\begin{aligned} z^* = \min \boldsymbol{c}^T\boldsymbol{x} \\ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ \boldsymbol{x} \in \mathbb{X} \end{aligned} \tag{1}$$

$$\begin{aligned} L(\boldsymbol{\mu}) = \min \boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\ \boldsymbol{x} \in \mathbb{X} \end{aligned} \tag{2}$$

Własność 2 (Zasada ograniczania (bounding principle))**.** *Dla dowolnego wektora (mnożników) $\boldsymbol{\mu}$, wartość $L(\boldsymbol{\mu})$ jest dolnym ograniczeniem wartości funkcji celu z^* problemu (1).*

Dowód. Rozwiązanie dopuszczalne (1) spełnia $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Zatem dla dowolnego $\boldsymbol{\mu}$

$$z^* = \min\{\boldsymbol{c}^T\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{X}\} = \min\{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{X}\}$$

Usuwając $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ nie zwiększymy funkcji celu, $z^* \geqslant \min\{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \mid \boldsymbol{x} \in \mathbb{X}\} = L(\boldsymbol{\mu})$. □

Dla dowolnego $\boldsymbol{\mu}$, $L(\boldsymbol{\mu})$ jest dolnym ograniczeniem na wartość funkcji celu problemu (1). Aby otrzymać najlepsze dolne ograniczenie (im wyższe tym lepsze) musimy rozwiązać **Problem mnożników Lagrange'a**

$$L^* = \max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}).$$

Z własności 2 mamy

$$L^* \leqslant z^*.$$

Dla dowolnego wyboru $\boldsymbol{\mu}$ i dowolnego rozwiązania dopuszczalnego $\boldsymbol{x}$ dostajemy

$$L(\boldsymbol{\mu}) \leqslant L^* \leqslant z^* \leqslant \boldsymbol{c}^T\boldsymbol{x}.$$

Powyższy warunek pozwala ocenić jakość rozwiązania dopuszczalnego.

$$(\boldsymbol{c}^T\boldsymbol{x} - L(\boldsymbol{\mu}))/L(\boldsymbol{\mu}).$$

Własność 3 (Test optymalności)**.**

- *Niech $\boldsymbol{\mu}$ będzie wektorem mnożników, a $\boldsymbol{x}$ będzie rozwiązaniem dopuszczalnym problemu (1) spełniającym $L(\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x}$. Wtedy $L(\boldsymbol{\mu})$ jest optymalną wartością funkcji celu problemu mnożników Lagrange'a ($L(\boldsymbol{\mu}) = L^*$) oraz $\boldsymbol{x}$ jest rozwiązaniem optymalnym problemu (1).*
- *Dla pewnego $\boldsymbol{\mu}$ rozwiązanie $\boldsymbol{x}^*$ problemu (2) (relaksacja Lagrange'a) jest rozwiązaniem dopuszczalnym problemu (1), wtedy $\boldsymbol{x}^*$ jest rozwiązaniem optymalnym problemu (1) i $\boldsymbol{\mu}$ jest rozwiązaniem optymalnym problemu mnożników Lagrange'a.*

Ograniczenia nierównościowe

$$\begin{aligned} z^* = \min \boldsymbol{c}^T \boldsymbol{x} \\ \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} \\ \boldsymbol{x} \in \mathbb{X} \end{aligned} \tag{3}$$

$$\begin{aligned} L(\boldsymbol{\mu}) = \min \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\ \boldsymbol{x} \in \mathbb{X} \end{aligned} \tag{4}$$

W tym przypadku **Problem mnożników Lagrange'a** ma postać

$$L^* = \max_{\boldsymbol{\mu} \geqslant \boldsymbol{0}} L(\boldsymbol{\mu}).$$

Może jednak się zdażyć, że *drugi warunek testu optymalności nie będzie spełniony* ($\boldsymbol{x}^*$ optymalny dla (4), dopuszczalny dla (3), nie optymalny dla (3)).

Rozwiązanie musi spełniać **warunek komplementarności**

$$\boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = 0.$$

Podsumowanie

- Każda wartość $L(\boldsymbol{\mu})$ jest dolnym ograniczeniem optymalnej wartości funkcji celu z^*.
- L^* jest najlepszym dolnym ograniczeniem.
- Jeżeli $L(\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x}$, to $L^* = z^* = \boldsymbol{c}^T\boldsymbol{x}$.
- L^* jest użytecznym w heurystykach i w metodzie podziału i ograniczeń.

2 Przykłady

Uogólnione zagadnienie przydziału (Generalized assignment problem)

Problem polega na przydziale $|I|$ "obiektów" do $|J|$ "pudełek". Przydzielamy obiekt dokładnie do jednego pudełka. Jeżeli przydzieliliśmy obiekt i do pudelka j wówczas obiekt i zużywa a_{ij} jednostek danego "zasobu" w tym pudelku oraz koszt przydziału równy jest c_{ij}. Łączna wielkość dostępnego zasobu pudełka j jest równa d_j. Szukamy dopuszczalnego przydziału, którego łączny koszt jest minimalny.

Przykład Szeregowanie zadań – obiektami są zadania, pudełkami są maszyny; a_{ij} jest czasem wykonania zadania i na maszynie j, a d_j jest łącznym czasem dostępności maszyny j.

$$\min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$$

$$\sum_{j\in J} x_{ij} = 1 \text{ dla } i \in I \tag{5}$$

$$\sum_{i\in I} a_{ij}x_{ij} \leqslant d_j \text{ dla } j \in J \tag{6}$$

$$x_{ij} \in \{0,1\} \quad i \in I, j \in J. \tag{7}$$

$x_{ij} = 1$ jeżeli przydzielamy obiekt i do pudelka j; $x_{ij} = 0$ w przeciwnym przypadku.

Możliwe są dwie relaksacje:

- Dodajemy do funkcji celu ograniczenie (5).
- Dodajemy do funkcji celu ograniczenie (6).

Dodajemy do funkcji celu ograniczenie (5)

$$L(\boldsymbol{\mu}) = \min \sum_{i\in I}\sum_{j\in J}(\mathbf{c_{ij}} + \mu_{\mathbf{i}})x_{ij} - \sum_{i\in I}\mu_i$$

$$\sum_{i\in I} a_{ij}x_{ij} \leqslant d_j \text{ dla } j \in J$$
$$x_{ij} \in \{0,1\} \quad i \in I, j \in J.$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{|I|})$$

Rozwiązanie powyższego zadania dla ustalonego $\boldsymbol{\mu}$ sprowadza się do rozwiązania $|J|$ niezależnych zagadnień plecakowych.

Dodajemy do funkcji celu ograniczenie (6)

$$L(\boldsymbol{\mu}) = \min \sum_{i\in J}\sum_{j\in I}(\mathbf{c_{ij}} + \mu_{\mathbf{j}}\mathbf{a_{ij}})x_{ij} - \sum_{i\in J}\mu_j d_j$$
$$\sum_{j\in J} x_{ij} = 1 \text{ dla } i \in I$$
$$x_{ij} \in \{0,1\} \quad i \in I, j \in J.$$
$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{|J|})$$

Co sprowadza się do rozwiązanie, dla ustalonego $\boldsymbol{\mu} \geqslant \mathbf{0}$, $|I|$ niezależnych łatwych problemów, $i \in I$,

$$\min \sum_{i\in J}(c_{ij} + \mu_j a_{ij})x_{ij}$$
$$\sum_{j\in J} x_{ij} = 1$$

Problem lokalizacji (Facility location problem)

Model dla problemu jest następujący

$$\min \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} + \sum_{j\in J} F_j y_j$$

$$\sum_{j\in J} x_{ij} = 1 \text{ dla } i \in I \tag{8}$$

$$\sum_{i\in I} d_i x_{ij} \leqslant K_j y_j \text{ dla } j \in J \tag{9}$$

$$0 \leqslant x_{ij} \leqslant 1 \text{ dla } i \in I \text{ dla } j \in J \tag{10}$$

$$y_j \in \{0,1\} \text{ dla } j \in J \tag{11}$$

- I zbiór odbiorców (klientów),
- J zbiór potencjalnych lokalizacji dostawców (np. hurtowni), z których zaopatruje się odbiorców,
- $y_j = 1$ jeżeli lokalizacja dostawcy jest w miejscu j, 0 w przeciwnym przypadku,
- wartość zmiennej $0 \leqslant x_{ij} \leqslant 1$ oznacza ułamek (procent) zapotrzebowania odbiorcy i pokrywany przez dostawcę w miejscu j,
- d_i jest zapotrzebowaniem odbiorcy i,
- c_{ij} są kosztami transportu między dostawcą zlokalizowanym w j, a odbiorca i,
- F_j jest stałym kosztem lokalizacji dostawcy (np. kosztem otwarcia, wynajęcia) w miejscu j,
- K_j jest podażą dostawcy w miejscu j.

Dodajemy do funkcji celu ograniczenie (8)

$$L(\boldsymbol{\mu}) = \min \sum_{i\in I}\sum_{j\in J} (\mathbf{c_{ij}} + \mu_{\mathbf{i}}) x_{ij} + \sum_{j\in J} F_j y_j - \sum_{i\in I} \mu_i$$

$$\sum_{i\in I} d_i x_{ij} \leqslant K_j y_j \text{ dla } j \in J$$

$$0 \leqslant x_{ij} \leqslant 1 \text{ dla } i \in I \text{ dla } j \in J$$

$$y_j \in \{0,1\} \text{ dla } j \in J$$

Co sprowadza się do rozwiązanie, dla ustalonego $\boldsymbol{\mu}$, $|J|$ niezależnych podproblemów (składający się z dwóch podproblemów **LP**), $j \in J$,

$$\begin{aligned} \min \sum_{i \in I} \sum_{j \in J} (\mathbf{c_{ij}} + \mu_{\mathbf{i}}) x_{ij} + \sum_{j \in J} F_j y_j \\ \sum_{i \in I} d_i x_{ij} \leqslant K_j y_j \\ 0 \leqslant x_{ij} \leqslant 1 \text{ dla } i \in I \\ y_j = 0 \text{ lub } y_j = 1 \end{aligned}$$

Dodajemy do funkcji celu ograniczenie (9)

$$L(\boldsymbol{\mu}) = \min \sum_{i \in I} \sum_{j \in J} (\mathbf{c_{ij}} + \mathbf{d_i} \mu_{\mathbf{j}}) x_{ij} + \sum_{j \in J} (F_j - \mu_j K_j) y_j$$

$$\begin{aligned} \sum_{j \in J} x_{ij} = 1 \text{ dla } i \in I \\ 0 \leqslant x_{ij} \leqslant 1 \text{ dla } i \in I \text{ dla } j \in J \\ y_j \in \{0, 1\} \text{ dla } j \in J \end{aligned}$$

Rozwiązujemy problem dla ustalonego $\boldsymbol{\mu} \geqslant \mathbf{0}$. $y_j = 1$ jeżeli $F_j - \mu_j K_j \leqslant 0$; 0 w przeciwnym przypadku. Po ustaleniu y_j, dla każdego $i \in I$ rozwiązujemy następujący łatwy problem (dla $j \in J$ ustalamy $x_{ij} = 1$)

$$\begin{aligned} \min \sum_{i \in I} \sum_{j \in J} (\mathbf{c_{ij}} + \mathbf{d_i} \mu_{\mathbf{j}}) x_{ij} \\ \sum_{j \in J} x_{ij} = 1 \\ 0 \leqslant x_{ij} \leqslant 1 \text{ dla } j \in I \end{aligned}$$

Uwagi na temat treści wykładu

Treść wykładu w całości została przygotowana na podstawie książki [1].

Literatura

[1] Ravindra K. Ahuja, Thomas L. Magnanti, James B. Orlin. *Network Flows: theory, algorithms, and applications.* Prentice Hall, Englewood Cliffs, New Jersey, 1993.