Zad 10.

wtorek, 28 marca 2023 11:53

10. (1pkt) Na wykładzie przedstawiono zachłanny algorytm dla problemu *Pokrycia zbioru*, znajdujący rozwiązania, które są co najwyżej $\log n$ razy gorsze od rozwiązania optymalnego.

    Pokaż, że istnieją dane, dla których rozwiązania znajdowane przez ten algorytm są blisko $\log n$ gorsze od rozwiązań optymalnych.

PROBLEM:

Dane: rodzina $S = \{S_1, S_2, \ldots, S_k\}$ podzbiorów n-elementowego uniwersum U
funkcja kosztu $c : S \to \mathcal{R}_+$

Zadanie: Znaleźć najtańszą podrodzinę S pokrywającą U, tj.
znaleźć $S' \subseteq S$ taką, że $\bigcup_{X \in S'} X = U$ i żadna inna podrodzina nie ma kosztu mniejszego od $Koszt(S')$, gdzie koszt podrodziny Z jest zdefiniowany jako

$$Koszt(Z) = \sum_{X \in Z} c(X).$$

(a) dla każdego podzbioru S_i określamy cenę za pokrywany element (w skrócie *cne*):

$$cne(S_i) = \frac{c(S_i)}{|S_i \setminus C|}$$

gdzie C oznacza zbiór dotychczas pokrytych elementów.

(b) do rozwiązania wybieramy ten z podzbiorów, dla którego wartość *cne* jest minimalna.

przykład takich danych

$U = \{a_1, a_2, a_3, \ldots, a_n\}$

$S = \{U, \{a_1\}, \{a_2\}, \ldots \{a_n\}\}$

$c : S \to \mathbb{R}_+$ t. że $c(S_i) = \frac{1}{i}$ $(i = 1 .. (n+1))$

$C_k \leftarrow$ zbiór pokrytych wierzchołków po k-tym kroku algorytmu

$k=0$

$C_0 = \emptyset$

$cne(U) = \frac{1}{n}$

$cne(\{a_1\}) = \frac{\frac{1}{2}}{1} = \frac{1}{2}$

$\vdots$

$cne(\{a_n\}) = \frac{\frac{1}{n+1}}{1} = \frac{1}{n+1}$ $\leftarrow$ najmniejsze cne

$k=1$ $C_1 = \{a_n\}$

$cne(U) = \frac{1}{n-1}$

$cne(\{a_1\}) = \frac{\frac{1}{2}}{1} = \frac{1}{2}$

$\vdots$

$\frac{1}{n}$ 1 znowu najmniejsze

$cne(\{\ldots\}) = \frac{\ldots}{1} = \ldots 2$

$\vdots$

$cne(\{a_{n-1}\}) = \frac{\frac{1}{n}}{1} = \frac{1}{n}$ ← znowu najmniejsze cne

$k=2$

$C_2 = \{a_n, a_{n-1}\}$ i tak dalej.

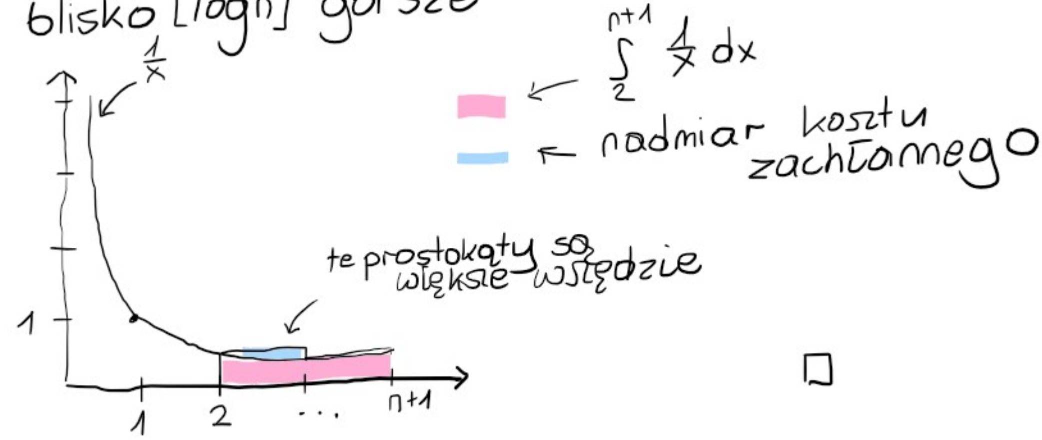
widać, że algorytm zachłanny wybierze wszystkie singletony z S_i

zapłaci $\sum_{k=2}^{n+1} \frac{1}{k}$. optymalne rozwiązanie ma jednak koszt 1 (bierzemy same U).
$= \Theta(1)$

$$\Theta(\log n) = \int_2^{n+1} \frac{1}{x}\, dx \leq \sum_{k=2}^{n} \frac{1}{k} \quad (*)$$

skoro koszt optymalnego rozwiązania jest rzędu $\Theta(1)$, a koszt zachłannego jest rzędu $\Theta(\log n)$, to jest blisko $[\log n]$ gorsze

(*)



□