Lista 3
Zadanie 2

Krystyna Korzonek

21 kwietnia 2020

1 Analiza algorytmu

Intuicja jest taka, że jak dzielimy zbiór S na równe podzbiory i drzewo wywołań funkcji jest niezbyt głębokie, to algorytm wykonuje mało porównań. Faktycznie okazuje się, że im bliższe potędze dwójki jest $|S| = n$, tym mniej dodatkowych porównań wykonuje algorytm. Dowód będzie dość techniczny i żmudny:
Niech f(n) będzie liczbą porównań wykonaną przez algorytm gdy $|S| = n$. Wprowadźmy oznaczenie:

$$\delta(n) = f(n) - \left\lceil \frac{3}{2}n - 2 \right\rceil$$

Rekurencyjny wzór na f(n) przedstawia się następująco:

$$f(n) = \begin{cases} 0 & gdy \quad n = 1 \\ 1 & gdy \quad n = 2 \\ f(\left\lceil \frac{n}{2} \right\rceil) + f(\left\lfloor \frac{n}{2} \right\rfloor) + 2 & w\,pp. \end{cases}$$

Pokażmy teraz następujące twierdzenie:

tw.: Jeśli $n \in [2^k, 2^{k+1}]$ i $r \in [0, 2^{k-1}]$, to $\delta(2^k + r) = \delta(2^{k+1} - r) = \left\lfloor \frac{r}{2} \right\rfloor$
Dowód będzie indukcyjny.
Podstawa indukcji $k = 1$

Rozpatrujemy $n \in \{2, 3, 4\}$, $r \leq 1$, $\left\lfloor \frac{r}{2} \right\rfloor = 0$. Chcemy pokazać, że $\delta(n) = 0$. Rzeczywiście:

$$f(2) = 1 = \left\lceil \frac{3}{2}2 - 2 \right\rceil$$

$$f(3) = 3 = \left\lceil \frac{3}{2}3 - 2 \right\rceil$$

$$f(4) = 4 = \left\lceil \frac{3}{2}4 - 2 \right\rceil$$

Krok indukcyjny tw(k-1) $\Rightarrow$ tw(k)
Przeprowadzę dowód dla $2^k + r$ (dla $2^{k+1} - r$ jest analogiczny):

$$f(2^k + r) = f\left(\left\lceil \frac{2^k + r}{2} \right\rceil\right) + f\left(\left\lfloor \frac{2^k + r}{2} \right\rfloor\right) + 2 =$$

$$= f\left(2^{k-1} + \left\lceil \frac{r}{2} \right\rceil\right) + f\left(2^{k-1} + \left\lfloor \frac{r}{2} \right\rfloor\right) + 2 =$$

Skorzystaliśmy z faktu, że liczbę całkowitą można wyciągnąć przed sufit/podłogę. Teraz zauważmy, że:

$$\left\lfloor \frac{r}{2} \right\rfloor \leq \left\lceil \frac{r}{2} \right\rceil \leq 2^{k-2}$$

(druga nierówność jest prawdziwa dzięki temu, że górne ograniczenie na r jest parzyste)
a więc możemy skorzystać z założenia indukcyjnego:

$$f(2^k+r) = \left\lceil \frac{3}{2}\left(2^{k-1} + \left\lceil \frac{r}{2} \right\rceil\right) - 2 \right\rceil + \left\lfloor \left\lceil \frac{r}{2} \right\rceil \frac{1}{2} \right\rfloor + \left\lceil \frac{3}{2}\left(2^{k-1} + \left\lfloor \frac{r}{2} \right\rfloor\right) - 2 \right\rceil + \left\lfloor \left\lfloor \frac{r}{2} \right\rfloor \frac{1}{2} \right\rfloor + 2 =$$

Wyciągamy części całkowite spod sufitów:

$$f(2^k + r) = 3 * 2^{k-1} - 2 + \left\lceil \frac{3}{2}\left\lceil \frac{r}{2} \right\rceil \right\rceil + \left\lfloor \left\lceil \frac{r}{2} \right\rceil \frac{1}{2} \right\rfloor + \left\lceil \frac{3}{2}\left\lfloor \frac{r}{2} \right\rfloor \right\rceil + \left\lfloor \left\lfloor \frac{r}{2} \right\rfloor \frac{1}{2} \right\rfloor$$

Przypomnijmy, że chcemy pokazać, że $\delta(2^k + r) = \left\lfloor \frac{r}{2} \right\rfloor$. Powyżej trochę "posprzątaliśmy"$f(2^k + r)$, teraz to wykorzystamy:

$$\delta(2^k + r) = f(2^k + r) - \left\lceil \frac{3}{2}\left(2^k + r\right) - 2 \right\rceil = f(2^k + r) - (3 * 2^{k-1} - 2) - \left\lceil \frac{3}{2}r \right\rceil =$$

$$= \left\lceil \frac{3}{2}\left\lceil \frac{r}{2} \right\rceil \right\rceil + \left\lfloor \left\lceil \frac{r}{2} \right\rceil \frac{1}{2} \right\rfloor + \left\lceil \frac{3}{2}\left\lfloor \frac{r}{2} \right\rfloor \right\rceil + \left\lfloor \left\lfloor \frac{r}{2} \right\rfloor \frac{1}{2} \right\rfloor - \left\lceil \frac{3}{2}r \right\rceil$$

Wobec tego żeby udowodnić tezę, wystarczy pokazać, że:

$$\left\lceil \frac{3}{2}\left\lceil \frac{r}{2} \right\rceil \right\rceil + \left\lfloor \left\lceil \frac{r}{2} \right\rceil \frac{1}{2} \right\rfloor + \left\lceil \frac{3}{2}\left\lfloor \frac{r}{2} \right\rfloor \right\rceil + \left\lfloor \left\lfloor \frac{r}{2} \right\rfloor \frac{1}{2} \right\rfloor = \left\lceil \frac{3}{2}r \right\rceil + \left\lfloor \frac{r}{2} \right\rfloor$$

Albo że:

$$(1) \qquad \left\lceil \frac{3}{2}\left\lceil \frac{r}{2} \right\rceil \right\rceil + \left\lceil \frac{3}{2}\left\lfloor \frac{r}{2} \right\rfloor \right\rceil = \left\lceil \frac{3}{2}r \right\rceil$$

$$(2) \qquad \left\lfloor \left\lceil \frac{r}{2} \right\rceil \frac{1}{2} \right\rfloor + \left\lfloor \left\lfloor \frac{r}{2} \right\rfloor \frac{1}{2} \right\rfloor = \left\lfloor \frac{r}{2} \right\rfloor$$

Najprościej można to pokazać rozpatrując r mod 4. Ponieważ to są bardzo proste i zupełnie nieciekawe przejścia, rozpiszę tylko jeden przypadek:

$$\left\lceil \frac{3}{2}\left\lceil \frac{4l+1}{2} \right\rceil \right\rceil + \left\lceil \frac{3}{2}\left\lfloor \frac{4l+1}{2} \right\rfloor \right\rceil = \left\lceil \frac{3}{2}(2l+2) \right\rceil + \left\lceil \frac{3}{2}2l \right\rceil = 6l+3 = \left\lceil \frac{3}{2}(4l+1) \right\rceil$$

Z prawdziwości (1) i (2) wynika teza.

Wnioski

1. Algorytm wykonuje minimalną liczbę porównań dla n takich, że $\delta(n) = 0$. Czyli dla n postaci $2^k + \epsilon$, $\epsilon \in \{-1, 0, 1\}$
2. Różnica liczby porównań wykonywanych przez algorytm od dolnej granicy dla $n \in [2^k, 2^{k+1}]$ jest równa co najwyżej $\left\lfloor \frac{r_{max}}{2} \right\rfloor = 2^{k-2}$

2 Ulepszenie algorytmu

Można zauważyć, że dążymy do jak najbardziej źrównoważonego podziału", więc dobrze byłoby dzielić zbiór S na podzbiór o parzystych mocach. Jeśli n jest nieparzyste i tak jeden podzbiór musi mieć nieparzyście wiele elementów, ale jeśli n jest postaci 4k + 2 można podzielić S na zbiory 2k i 2k+2-elementowe.
Oczywiście warunku nie możemy sformułować wprost zależnie od n mod 4, bo w tym byłoby ukryte porównanie. Potrzebujemy jakiegoś wyrażenia, które zwróci mniejszą z mocy podzbiorów na które dzielimy, tzn. które będzie spełniało:

$$w(n) = \begin{cases} 2k & gdy \quad n = 4k \\ 2k & gdy \quad n = 4k+1 \\ 2k & gdy \quad n = 4k+2 \\ 2k+1 & gdy \quad n = 4k+3 \end{cases}$$

Może to być np.:

$$w(n) = 2\left\lfloor \frac{n}{4} \right\rfloor + \left\lfloor \frac{n \bmod 4}{3} \right\rfloor$$

Teraz pokażmy, że g(n) - liczba porównań wykonywana przez poprawiony algorytm - jest równa dolnej granicy. Dowód będzie indukcyjny.

Podstawa indukcji

$$g(1) = 0 = \left\lceil \frac{3}{2} - 2 \right\rceil$$

$$g(2) = 1 = \left\lceil \frac{3}{2}2 - 2 \right\rceil$$

Założenie indukcjne

dla $k < n$ zachodzi $g(k) = \left\lceil \frac{3}{2}k - 2 \right\rceil$

Krok indukcyjny

Załóżmy, że algorytm dzieli n w sposób opisany powyżej na n_1 i n_2, n_1 i n_2 są oczywiście mniejsze od n, korzystamy z założenia indukcyjnego:

$$g(n) = g(n_1) + g(n_2) + 2 = \left\lceil \frac{3}{2}n_1 - 2 \right\rceil + \left\lceil \frac{3}{2}n_2 - 2 \right\rceil + 2$$

Jeśli n_1 i n_2 są parzyste, to:

$$g(n) = \frac{3}{2}(n_1 + n_2) - 2 = \frac{3}{2}n - 2 = \left\lceil \frac{3}{2}n - 2 \right\rceil$$

Jeśli n_1 jest parzyste, a n_2 nieparzyste, skorzystamy z następującego faktu:

fakt jeśli m jest nieparzyste, to $\left\lceil \frac{3}{2}m - 2 \right\rceil = \frac{3}{2}m - \frac{3}{2}$

$$g(n) = \left\lceil \frac{3}{2}n_1 - 2 \right\rceil + \left\lceil \frac{3}{2}n_2 - 2 \right\rceil + 2 = \frac{3}{2}n_1 + \left\lceil \frac{3}{2}n_2 - 2 \right\rceil = \frac{3}{2}(n_1 + n_2) - \frac{3}{2} =$$

$$= \left\lceil \frac{3}{2}n - 2 \right\rceil$$

n_1 i n_2 nie mogą być jednocześnie nieparzyste, rozpatrzyliśmy więc wszystkie przypadki. Opisany sposób wykonuje więc zawsze najmniejszą możliwą liczbę porównań.