Lista 5, zadanie 4 - Tomasz Woszczyński

Treść: Udowodnij, że $2n-1$ porównań trzeba wykonać, aby scalić dwa ciągi n elementowe w modelu drzew decyzyjnych. Zastosuj grę z adwersarzem, w której adwersarz na początku ogranicza przestrzeń danych tak, by zawierała $2n$ zestawów danych takich i by każde porównanie wykonane przez algorytm eliminowało co najwyżej jeden zestaw.

Rozwiązanie: Mamy dwa n elementowe ciągi, nazwijmy je X oraz Y, chcemy z nich utworzyć posortowany ciąg S. Wiemy, że $|S| = 2n$. Gdybyśmy nie zwracali uwagi na to, że wejściowe ciągi są posortowane, to mielibyśmy aż $\binom{2n}{n}$ możliwych ułożeń wszystkich elementów, jednak taka przestrzeń danych jest zbyt duża i adwersarz będzie chciał się pozbyć jak największej ilości elementów, które może łatwo wykluczyć, aby gracz musiał wykonać jak najwięcej porównań, by znaleźć odpowiedź. Mając informację o tym, iż ciągi X, Y są posortowane, możemy stworzyć własną przestrzeń zdarzeń S_i, tworząc każdy kolejny ciąg poprzez zamianę elementów s_k oraz s_{k+1} (dla $k > 0$). Pierwszym, bazowym ciągiem S_0 będzie ciąg:

$$x_1, y_1, x_2, y_2, \dots, x_n, y_n \equiv s_1, s_2, \dots s_{2n}$$

Wspomnianych zamian w $2n$ elementowym ciągu możemy wykonać $2n-1$, a więc tych ciągów będzie $2n$. Kolejne ciągi S_i będą wyglądać następująco:

$$
\begin{gathered}
S_0 = x_1, y_1, x_2, y_2, x_3, y_3, \dots, x_n, y_n \\
S_1 = y_1, x_1, x_2, y_2, x_3, y_3, \dots, x_n, y_n \\
S_2 = x_1, x_2, y_1, y_2, x_3, y_3, \dots, x_n, y_n \\
S_3 = x_1, y_1, y_2, x_2, x_3, y_3, \dots, x_n, y_n \\
S_4 = x_1, y_1, x_2, x_3, y_2, y_3, \dots, x_n, y_n \\
S_5 = x_1, y_1, x_2, y_2, y_3, x_3, \dots, x_n, y_n \\
\vdots \\
S_{2n-1} = x_1, y_1, x_2, y_2, y_3, x_3, \dots, y_n, x_n
\end{gathered}
$$

Mając tak przygotowane ciągi, możemy rozpatrzeć cztery możliwe odpowiedzi na pytanie "jak ułożone powinny być wartość x_i oraz wartość y_j" (pomijamy pytania o ułożenie elementów o indeksach i, j w ramach jednego ciągu, gdyż są one posortowane i nie mają w tym przypadku żadnego sensu):

1. $i = j$: w takiej sytuacji otrzymujemy odpowiedź, że $x_i < y_j$, dzięki czemu eliminujemy przypadek $x_i > y_j$, a więc zamianę elementów $s_{2i-1} = x_i$ oraz $s_{2i} = y_i$ w S_0, a więc w ciągu S_{2i-1}.

2. $i = j + 1$: w takiej sytuacji otrzymujemy odpowiedź, że $x_i > y_j$, dzięki czemu eliminujemy przypadek, w którym $x_i < y_{i-1}$, a więc zamianę elementów $s_{2i-2} = y_{i-1}$ oraz $s_{2i-1} = x_i$ w S_0, czyli ciąg S_{2i-2}.

3. $i > j + 1$: w takiej sytuacji otrzymujemy odpowiedź, że $x_i > y_i$ - tutaj nie eliminujemy żadnego ze swoich zestawów, gdyż kolejne zestawy to zamiana x_i z y_i oraz x_{i+1} z y_i, a elementy x_i oraz y_j różnią się o co najmniej 2.

4. $j > i$: w takiej sytuacji otrzymujemy odpowiedź, że $x_i < y_i$, jest to sytuacja analogiczna do przypadku trzeciego, nie eliminujemy więc żadnego zestawu, ponieważ w kolejnych elementach s_k nie zamieniamy ze sobą elementów x_i i y_j dla $i < j$, a w S_0 dla $i < j$ zachodzi $x_i < y_j$.

Dzięki rozpatrzeniu powyższych możliwości wiemy, że dla dowolnego z zapytań usuniemy maksymalnie jeden zestaw. Wiedząc, że adwersarz chce utrudnić nam życie i sprowokować do zadania jak największej ilości pytań. W tym przypadku przestrzeń wszystkich ciągów miała moc $2n$ i każde pytanie eliminowało co najwyżej jeden zestaw, co oznacza, że należy zadać co najmniej $2n - 1$ pytań (lub więcej, gdy są one "głupie", a więc dotyczą informacji, które już mamy), by znaleźć odpowiedź. Skoro nie możemy eliminować więcej niż jednego zestawu na pytanie, nie możemy zadać mniejszej ilości pytań, oznacza to, że należy wykonać $2n - 1$ porównań, aby w modelu drzew decyzyjnych scalić dwa posortowane ciągi X, Y. $\blacksquare$