Gra z adwersarzem: zadanie 5.2

Kamil Matuszewski

10 maja 2016

Treść

Udowodnij, że $2n-1$ porównań trzeba wykonać, aby scalić dwa ciągi n elementowe w modelu drzew decyzyjnych. Użyj gry z adwersarzem, który najpierw ogranicza przestrzeń danych do $2n$ tak, by każde porównanie eliminowało co najwyżej jeden zestaw.

Przestrzeń danych i jej ograniczenie

Mamy więc dwa ciągi n elementowe: $A = a_1, a_2, a_3, \ldots, a_n$ i $B = b_1, b_2, b_3, \ldots, b_m$, i chcemy z nich utworzyć posortowany ciąg X, złożony ze wszystkich elementów z A i wszystkich elementów z B (więc jest długości $2n$). Adwersarz na początku przygotowuje sobie możliwe odpowiedzi - przestrzeń danych. Wszystkich możliwych ustawień ciągu złożonego z elementów z A i z B jest $\binom{2n}{n}$ - wybieramy miejsca dla elementów np. z ciągu A z $2n$ dostępnych. Skoro A i B są posortowane (inaczej scalanie nie ma sensu) to wybranie miejsc dla któregoś z ciągów jednoznacznie daje nam ułożenie ciągu X. Adwersarz chce jednak jak najbardziej ograniczyć przestrzeń danych, by każde zapytanie eliminowało jak najmniej elementów a przestrzeń zdarzeń była jak największa. Innymi słowy chce, żeby gracz musiał wykonać jak najwięcej porównań by znaleźć odpowiedź.
Ograniczmy więc przestrzeń zdarzeń, tworząc ciągi, które wyglądają następująco: pierwszy ciąg to $X_0 = a_1, b_1, a_2, b_2, \ldots, a_n, b_n = x_1, x_2, x_3, \ldots, x_{2n}$ (oczywiście jest on posortowany rosnąco), a każdy kolejny X_k dla $k > 0$ tworzymy zamieniając ze sobą elementy x_k i x_{k+1}. Takich zamian w ciągu o $2n$ elementach możemy wykonać $2n-1$. W sumie mamy więc $2n$ ciągów. To będzie nasza przestrzeń zdarzeń.
Dla jasności:

$$X_0 = a_1, b_1, a_2, b_2, \ldots, a_n, b_n$$

$$X_1 = b_1, a_1, a_2, b_2, \ldots, a_n, b_n$$

$$X_2 = a_1, a_2, b_1, b_2, \ldots, a_n, b_n$$

$$X_3 = a_1, b_1, b_2, a_2, \ldots, a_n, b_n$$

$$\vdots$$

$$X_{2n-1} = a_1, b_1, a_2, b_2, \ldots, b_n, a_n$$

Zapytania gracza

Po pierwsze załóżmy, że gracz nie zadaje głupich pytań, w stylu "Czy a_i jest większe od a_j", gdyż ciągi A i B są posortowane, i takie pytania nic by nie wniosły (a gracz pragnie szybko wygrać). Weźmy dowolne zapytanie "jakie jest a_i w stosunku do b_j". Rozpatrzmy trzy przypadki:

- $i > j + 1$. Adwersarz odpowiada wtedy a_i jest większe od b_j. Zauważmy, że w ten sposób nie eliminuje żadnego ze swoich zestawów - kolejne zestawy to zamiana b_i z a_i oraz b_i z a_{i+1}, a a_i i b_j różnią się co najmniej o 2.
- $j > i$. Podobne do przypadku powyżej. Adwersarz odpowiada, że a_i jest mniejsze od b_j. To pytanie również nie eliminuje żadnego z zestawów adwersarza, bowiem w kolejnych x_k nie zamieniamy ze sobą elementów a_i i b_j dla $i < j$, a w x_0 dla $i < j$ zachodzi $a_i < b_j$.
- $i = j$. Wtedy adwersarz odpowiada, że a_i jest mniejsze od b_j. Zauważmy, że w ten sposób eliminuje przypadek, w którym a_i jest większe od b_i - jest to zamiana $x_{2i-1} = a_i$ i $x_{2i} = b_i$ w ciągu X_0, czyli ciąg X_{2i-1}.
- $i = j + 1$. Wtedy adwersarz odpowiada, że a_i jest większe od b_j. Zauważmy, że w ten sposób eliminuje przypadek, w którym a_i jest mniejsze od b_{i-1} - jest to zamiana $x_{2i-2} = b_{i-1}$ i $x_{2i-1} = a_i$ w ciągu X_0, czyli ciąg X_{2i-2}.

Pokazaliśmy więc, że dla dowolnego zapytania, usuniemy co najwyżej jeden zestaw.

Podsumowanie

Gra z adwersarzem opiera się na prostych założeniach. Adwersarz jest troszkę niemiły, i na początku nie zna odpowiedzi na pytanie, choć twierdzi, że ją zna. Nie chce bowiem, by gracz zbyt szybko odgadł. Pragnie zmusić go do zadania jak największej liczy pytań. W tym celu najpierw tworzy sobie zestaw możliwych odpowiedzi tak, by było ich jak najwięcej, i każde pytanie eliminowało jak najmniej zestawów. W naszym przypadku adwersarz ograniczył sobie przestrzeń do $2n$, i każde jego pytanie eliminowało co najwyżej jeden z tych zestawów. Oznacza to, że potrzeba przynajmniej $2n - 1$ pytań, by znaleźć odpowiedź w tym zestawie. Gdyby gracz miał strategię, która pozwala znaleźć odpowiedź wcześniej, w szczególności znalazłby ją wcześniej w tym zestawie, eliminując więcej niż jedną odpowiedź, co jak pokazaliśmy jest niemożliwe. To oznacza, że potrzeba $2n - 1$ porównań, by w modelu drzew decyzyjnych scalić dwa ciągi.