Wojciech Rytter — Treść zadania, Opracowanie
Arkadiusz Paterek — Program

OI, Etap III, dzień 1, 6–04–2005

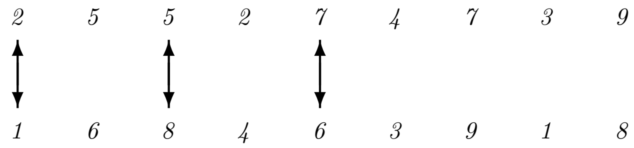
Dwuszereg

W dwuszeregu stoi $2n$ żołnierzy. Trzeba przestawić żołnierzy tak, aby w tym samym szeregu nie było dwóch żołnierzy tego samego wzrostu — wówczas powiemy, że żołnierze są ustawieni poprawnie.

Pojedyncza operacja polega na zamianie miejscami dwóch żołnierzy, którzy są na tej samej pozycji w obu szeregach. Twoim zadaniem jest policzenie minimalnej liczby zamian, jakie trzeba wykonać, aby żołnierze byli ustawieni poprawnie.

Przykład

Na rysunku mamy dwuszereg złożony z 18 żołnierzy. Strzałkami zaznaczono 3 operacje zamiany, po wykonaniu których żołnierze są ustawieni poprawnie.



Zadanie

Napisz program, który:

- *wczyta ze standardowego wejścia liczbę i wzrost żołnierzy, tak jak są ustawieni na początku,*
- *wyznaczy minimalną liczbę zamian miejscami (żołnierzy stojących na tej samej pozycji w obu szeregach) potrzebnych do poprawnego ustawienia żołnierzy,*
- *wypisze wynik na standardowe wyjście.*

Wejście

W pierwszym wierszu wejścia znajduje się jedna liczba całkowita n, $1 \leqslant n \leqslant 50\,000$. W każdym z dwóch szeregów stoi n żołnierzy. W każdym z dwóch kolejnych wierszy znajduje się po n dodatnich liczb całkowitych pooddzielanych pojedynczymi odstępami. W drugim wierszu znajdują się liczby $x_1, x_2, \ldots, x_n$, $1 \leqslant x_i \leqslant 100\,000$; x_i to wzrost i-go żołnierza w pierwszym szeregu. W trzecim wierszu znajdują się liczby $y_1, y_2, \ldots, y_n$, $1 \leqslant y_i \leqslant 100\,000$; y_i to wzrost i-go żołnierza w drugim szeregu.

Możesz założyć, że dla danych testowych zawsze możliwe jest poprawne ustawienie żołnierzy.

Wyjście

W pierwszym i jedynym wierszu wyjścia powinna znaleźć się jedna nieujemna liczba całkowita — minimalna liczba zamian jakie należy wykonać, aby żołnierze byli poprawnie ustawieni.

Przykład

Dla danych wejściowych:

```
9
2 5 5 2 7 4 7 3 9
1 6 8 4 6 3 9 1 8
```

poprawnym wynikiem jest:

```
3
```

Rozwiązanie

Wstępne uwagi i oznaczenia

Przyjmiemy następujące oznaczenia.

- Wysokości żołnierzy z pierwszego szeregu są zapisane w tablicy $\mathcal{S}_1[1..n]$, a wysokości żołnierzy z drugiego szeregu — w tablicy $\mathcal{S}_2[1..n]$.
- Pozycje $\mathcal{S}_1[i]$ i $\mathcal{S}_2[i]$ nazwiemy *kolumną i*.
- Operacja $zamiana(i)$ powoduje zamianę miejscami żołnierzy stojących w i-tej kolumnie: $\mathcal{S}_1[i] \leftrightarrow \mathcal{S}_2[i]$.
- Dwie różne kolumny i, j nazwiemy *zależnymi*, gdy $\{ \mathcal{S}_1[i], \mathcal{S}_2[i] \} \cap \{ \mathcal{S}_1[j], \mathcal{S}_2[j] \} \neq \emptyset$.

Zauważmy także, iż aby istniało rozwiązanie zadania (a taką gwarancję mamy w treści), nie może być *trzech żołnierzy jednakowego wzrostu*.

Interpretacja grafowa

Definicja 1 *„Graf" zależności* jest to graf $G = (V, E)$ o zbiorze wierzchołków $V = \{1, 2, \ldots, n\}$ i zbiorze krawędzi E łączących wierzchołki odpowiadające kolumnom zależnym. W przypadku kolumn i, j ($i \neq j$), dla których $\{ \mathcal{S}_1[i], \mathcal{S}_2[i] \} = \{ \mathcal{S}_1[j], \mathcal{S}_2[j] \}$, wierzchołki i oraz j są połączone dwiema krawędziami (z tego powodu słowo graf w niniejszej definicji zostało ujęte w cudzysłów).

Możemy zauważyć, że graf zależności ma bardzo prostą strukturę.

Lemat 1 *Graf zależności G składa się z rozłącznych cykli i ścieżek.*

Dowód Wystarczy zauważyć, że z wierzchołka odpowiadającego kolumnie i mogą wychodzić krawędzie tylko do kolumn, w których stoi żołnierz o wzroście $\mathcal{S}_1[i]$ lub $\mathcal{S}_2[i]$. Ponieważ jest najwyżej dwóch żołnierzy o wzroście $\mathcal{S}_1[i]$, z tego jeden w kolumnie i — podobnie dla $\mathcal{S}_2[i]$ — więc z wierzchołka i mogą wychodzić najwyżej dwie krawędzie. Stąd w grafie zależności każdy wierzchołek ma co najwyżej stopień dwa, czyli graf musi być złożony z rozłącznych cykli i ścieżek. ■

Zmodyfikujmy graf zależności nadając każdej krawędzi zwrot w ten sposób, by powstały skierowane cykle i skierowane ścieżki (patrz rysunek). Przypiszmy także krawędziom grafu etykiety, które nazwiemy *kolorami*. Dla krawędzi $e = i \to j$ określmy:

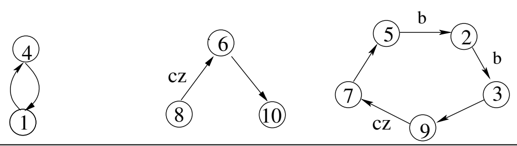
$$kolor(e) = \begin{cases} \textit{czarny} & \text{gdy } \mathcal{S}_1[i] = \mathcal{S}_2[j] \\ \textit{biały} & \text{gdy } \mathcal{S}_2[i] = \mathcal{S}_1[j] \end{cases}$$

W przypadku cyklu dwuelementowego (i, j), dla którego powstaje niejednoznaczność przy określaniu kolorów zgodnie z powyższym wzorem, umówmy się, że gdy $\mathcal{S}_2[i] = \mathcal{S}_1[j]$ (i tym samym $\mathcal{S}_1[i] = \mathcal{S}_2[j]$), to obu krawędziom cyklu przypisujemy kolor biały.

W pozostałych przypadkach będziemy mówić, że *krawędź nie ma koloru*.

Przykład.

numer kolumny	1	2	3	4	5	6	7	8	9	10
pierwszy szereg	2	6	5	2	7	4	7	3	9	4
drugi szereg	1	5	8	1	6	3	9	10	8	11



Rys. 1: Konfiguracja początkowa i odpowiadający jej skierowany graf zależności z etykietami krawędzi.

■ *Przykład*

Porządkowanie ustawienia

Rozważmy teraz jedną składową grafu zależności i zajmijmy się doprowadzeniem do poprawnego ustawienia żołnierzy z tej składowej (możemy to robić niezależne od „porządkowania” pozostałych składowych). Po pierwsze zauważmy, że prawdziwy jest następujący lemat:

Lemat 2 *Ustawienie żołnierzy z kolumn jednej składowej grafu zależności jest poprawne wtedy i tylko wtedy, gdy wszystkie krawędzie tej składowej mają taki sam kolor.*

Dowód Po pierwsze zauważmy, że przy poprawnym ustawieniu wszystkie krawędzie rozważanej składowej muszą mieć określony kolor — w przeciwnym razie występuje błąd ustawienia ($\mathcal{S}_1[i] = \mathcal{S}_1[j]$ lub $\mathcal{S}_2[i] = \mathcal{S}_2[j]$).

Rozważmy teraz przypadek, gdy składowa ma przynajmniej trzy wierzchołki. Wystarczy wówczas zauważyć, że jeśli krawędzie $i \to j$ oraz $j \to k$ są jej kolejnymi krawędziami i obie mają określony kolor, to musi to być ten sam kolor. W przeciwnym razie jedna wartość wzrostu żołnierza musiałaby wystąpić w danych przynajmniej trzykrotnie ($\mathcal{S}_1[i] = \mathcal{S}_2[j] = \mathcal{S}_1[k]$ lub $\mathcal{S}_2[i] = \mathcal{S}_1[j] = \mathcal{S}_2[k]$).

Pozostaje rozważyć przypadek, gdy składowa ma dwa wierzchołki. W takiej sytuacji albo składowa (ścieżka) ma dokładnie jedną krawędź, albo jest to cykl i wówczas również widać, że obie jego krawędzie muszą mieć ten sam kolor. ■

Z lematu wynika, że dla każdej składowej grafu zależności istnieją dokładnie dwa poprawne ustawienia. Pozostaje rozstrzygnąć, do którego z nich możemy doprowadzić wykonując mniejszą liczbę zamian. Dla składowej $\mathcal{X}$ i koloru x oznaczmy przez $LZ(\mathcal{X},x)$ minimalną liczbę zamian, które trzeba wykonać dla kolumn tworzących tę składową, by wszystkie jej krawędzie miały kolor x. Zakładając, że potrafimy znajdować kolejne wierzchołki składowej w czasie stałym, mamy następujący lemat.

Lemat 3 *Dla każdej składowej grafu zależności $\mathcal{X}$ i każdego koloru x wartość $LZ(\mathcal{X},x)$ można obliczyć w czasie liniowo zależnym od liczby wierzchołków składowej $\mathcal{X}$.*

Dowód W niniejszym dowodzie będziemy mówili, że krawędź $i \rightarrow j$ *zaczyna się w pierwszym szeregu*, gdy jest czarna lub $\mathcal{S}_1[i] = \mathcal{S}_1[j]$. W przeciwnym razie powiemy, że krawędź *zaczyna się w drugim szeregu*. Analogicznie powiemy, że krawędź $i \rightarrow j$ *kończy się w pierwszym szeregu*, gdy jest to krawędź biała lub $\mathcal{S}_1[i] = \mathcal{S}_1[j]$. W przeciwnym razie powiemy, że krawędź ta *kończy się w drugim szeregu*. Łatwo zauważyć, że składowa jest biała, gdy wszystkie jej krawędzie zaczynają się w drugim szeregu i kończą w pierwszym. Natomiast składowa jest czarna, gdy wszystkie jej krawędzie zaczynają się w pierwszym szeregu i kończą w drugim.

Rozważmy najpierw składową-ścieżkę $\mathcal{X} = (i_1, i_2, \ldots, i_k)$ i niech x będzie kolorem *białym*. Jeśli $\mathcal{S}_1[i_1] = \mathcal{S}_2[i_2]$ lub $\mathcal{S}_1[i_1] = \mathcal{S}_1[i_2]$ (czyli krawędź $i_1 \rightarrow i_2$ zaczyna się w pierwszym szeregu), to dokonujemy zamiany w kolumnie i_1. Następnie w kolejnych kolumnach i_j dla $j = 2, 3, \ldots, k$ dokonujemy zamiany tylko wówczas, gdy krawędź $i_{j-1} \rightarrow i_j$ kończy się w drugim szeregu. Łatwo zauważyć, że po wykonaniu wszystkich zamian uzyskamy składową w kolorze białym. Co więcej, zauważmy, że wszystkie dokonane zmiany **były konieczne**. Operacja $zamiana(i_1)$ była wykonywana tylko wówczas, gdy pierwsza krawędź zaczynała się w pierwszej kolumnie. Pozostałe zamiany wykonywaliśmy tylko wówczas, gdy rozważana krawędź $(i_{j-1} \rightarrow i_j)$ kończyła się w drugim szeregu. W ten sposób przeglądając jednokrotnie składową $\mathcal{X}$ wyznaczyliśmy wartość $LZ(\mathcal{X}, \textit{biały})$. Analogicznie (lub jednocześnie) możemy wyznaczyć wartość $LZ(\mathcal{X}, \textit{czarny})$.

Niech teraz $\mathcal{X} = (i_1, i_2, \ldots, i_k)$ będzie składową-cyklem i ponownie przyjmijmy najpierw, że x to kolor biały. Znajdujemy minimalne j, dla którego krawędź $(i_{j-1} \rightarrow i_j)$ kończy się w drugim szeregu (jeśli krawędź $i_k \rightarrow i_1$ kończy się w drugim szeregu, to przyjmujemy $j = 1$). Zauważmy, że znaleziona krawędź musi być czarna lub niepokolorowana. Dokonujemy zamiany w kolumnie i_j i poszukujemy kolejnej krawędzi cyklu kończącej się w drugim szeregu. Dokonujemy dla niej analogicznej zamiany jak poprzednio i kontynuujemy proces, aż wszystkie krawędzie będą kończyć się w pierwszym szeregu (tym samym wszystkie krawędzie cyklu będą zaczynać się w drugim szeregu, a więc wszystkie będą białe). Podobnie jak w przypadku składowej-ścieżki łatwo zauważyć, że wszystkie wykonane zamiany były konieczne, by osiągnąć białe pokolorowanie składowej. Także podobnie jak poprzednio, analogicznie możemy wyznaczyć wartość $LZ(\mathcal{X}, \textit{czarny})$.

■

Przykład c.d.

Dla naszego przykładu mamy dwie składowe-cykle: $X_1 = \{1,4\}$, $X_3 = \{2,3,9,7,5\}$ oraz jedną ścieżkę: $X_2 = \{8,6,10\}$. Początkowo krawędzie $8 \rightarrow 6$ i $9 \rightarrow 7$ mają kolor czarny, natomiast krawędzie $2 \rightarrow 3$ i $5 \rightarrow 2$ mają kolor biały.

Aby składowej X_1 nadać białe etykiety, wystarczy wykonać operację zamiany na pozycji 1 albo na pozycji 4. Zgodnie z wyjątkiem opisanym w definicji etykietowania, tej składowej nie da się pokolorować inaczej.

Aby składowej X_3 nadać białe etykiety, wystarczy wykonać zamiany na pozycjach: 7 i 9. By doprowadzić do pokolorowania składowej na czarno, trzeba wykonać zamiany w kolumnach 2, 3 i 5.

Pozostaje jeszcze składowa X_2. Pokolorowanie jej na biało uzyskamy dokonując zamian w kolumnach 6 i 8. Aby pokolorować ją na czarno, wystarczy zamienić miejscami żołnierzy w kolumnie 10.

Stąd optymalny sposób uzyskania poprawnego ustawienia wymaga 4 zamian.

numer kolumny	1	2	3	4	5	6	7	8	9	10
pierwszy szereg	2	5	8	1	6	4	7	3	9	11
drugi szereg	1	6	5	2	7	3	9	10	8	4
				*			*		*	*

Rys. 2: Ustawienie, w którym składowe X_1 i X_3 są pokolorowane na biało, a składowa X_2 na czarno. Gwiazdkami zaznaczono kolumny, w których należy wykonać zamiany, by uzyskać jednobarwne kolorowanie każdej składowej. Jest to optymalny sposób uporządkowania dwuszeregu.

■ *Przykład c.d.*

Algorytm

Podsumowując opisane rozważania możemy zapisać algorytm dla zadania.

```
procedure Algorytm
  begin
    LZ:=0;
    forall składowa X do
      LZ:=LZ+min{LZ(X,czarny),LZ(X,biały)};
    return LZ;
  end
```

Algorytm można zaimplementować tak, by działał w czasie $O(n)$. Zakładamy, że dane (wysokości żołnierzy) są na tyle małymi liczbami całkowitymi, że można je posortować w czasie liniowym (np. za pomocą algorytmu *radixsort*, jego opis można znaleźć w [12], [14] lub [17]). Wtedy możemy zastąpić je ich numerami w porządku, czyli kolejnymi liczbami naturalnymi. Pozwala to wyznaczyć graf zależności w czasie liniowym. Pozostałe operacje, czyli wyznaczenie składowych i etykiet ich krawędzi, wymagają już tylko jednokrotnego odwiedzenia każdego wierzchołka grafu.

Ciąg dalszy

Oprócz wyznaczenia poprawnego ustawienia osiąganego przez minimalną liczbę zamian, możemy także łatwo wyznaczyć liczbę poprawnych ustawień żołnierzy — wynosi ona 2^r, gdzie r jest liczbą składowych grafu zależności, włącznie ze składowymi jednoelementowymi (ale tylko tymi, dla których w odpowiadających im kolumnach stoją dwaj żołnierze różnego wzrostu).

Możemy także zmodyfikować zadanie przyjmując, że żołnierze stoją w trzech, a nie w dwóch szeregach. Jedna zamiana jest teraz jakąkolwiek wymianą trzech żołnierzy stojących w tej samej kolumnie. Opisany algorytm w tej sytuacji nie działa. Pozostawiamy Czytelnikowi zastanowienie się nad rozwiązaniem tak zmienionego zadania. Czy jest równie proste jak poprzednie?

Testy

Rozwiązanie było testowane na zestawie 15 testów. Opis poszczególnych testów znajduje się poniżej, *n* oznacza w nich liczbę żołnierzy stojących w szeregu.

Nazwa	**n**	**Opis**
dwu1.in	100	krótkie cykle
dwu2.in	200	średnie ścieżki
dwu3.in	500	średnie cykle
dwu4.in	1 000	długie ścieżki
dwu5.in	2 000	długie cykle
dwu6.in	5 000	krótkie ścieżki
dwu7.in	50 000	krótkie cykle, krótkie ścieżki
dwu8.in	50 000	krótkie cykle, średnie ścieżki
dwu9.in	50 000	krótkie cykle, długie ścieżki
dwu10.in	50 000	średnie cykle, krótkie ścieżki
dwu11.in	50 000	średnie cykle, średnie ścieżki
dwu12.in	50 000	średnie cykle, długie ścieżki
dwu13.in	50 000	długie cykle, krótkie ścieżki
dwu14.in	50 000	długie cykle, średnie ścieżki
dwu15.in	50 000	długie cykle, długie ścieżki