2019 DSA Mid-term Programming Exam

Things to remember before you get started...

- As usual, your program should take the input from the standard input and send the output to the standard output. That is, you should be able to redirect the input/output files like this:
  ./myProgram < input.txt > output.txt
- Time and memory limits for each problem are listed at the judge system.
- Some of the subtasks are relatively easy and you may want to use simple methods to get credits from these easy subtasks. (Please check out the specs of each subtask carefully.)
- In order to read integers faster, it is highly recommended to use the function rit() to read integers, as shown in the following example which prints the sum of two integers:

```
#include <iostream>
using namespace std;

void rit(int &x) {
    char c=0; bool fg=0;
    while (c=getchar(), (c<'0' && c!='-') || c>'9');
    c=='-' ? (fg=1,x=0) : (x=c-'0');
    while (c=getchar(), c>='0' && c<='9') x=x*10+(c-'0');
    if (fg) x=-x;
}

int main () {
    int a,b; rit(a); rit(b);
    cout << a+b << '\n';
}
```

And here goes the exam:

1. (30 pts) **Running order statistic**: Given a stream of N numbers coming one after another, we want to calculate the pth smallest number of the current received set. Specifically, if the ith number added to the set is denoted by a_i (starting from $i = 0$), then we want to compute b_i, which is the pth smallest number in the set, with $p = \lfloor \frac{i}{K} \rfloor$ and K being a given constant. (Note that the smallest number in the set is the 0th element.)

   After calculating the sequence of b_i, please print the value of the following expression to standard output: $(b_0 \times 1 + b_1 \times 2 + \cdots + b_{N-1} \times N)$ mod 2^{32}

   More details:

   - Input format:
     - Line 1 contains an integer T which denotes the number of test cases in the file.
       - For each test case, the first line contains two integers N,K, which are the number of integers you need to add into the set, and the parameter mentioned in the problem description, respectively.
       - In the next line, there are N space-separated integers $a_0, a_1, \ldots, a_{N-1}$, as mentioned in the problem description.
   - Output format:
     - A single number of the output
   - Sample test case
     - Input file

```
2
5 2
3 1 2 3 4
3 3
1 2 3
```

- Output file

```
34
6
```

- More explanation: In the first test case in sample input, $[b_0, b_1, b_2, b_3, b_4] = [3, 1, 2, 2, 3]$ so the answer is $b_0 \times 1 + b_1 \times 2 + b_2 \times 3 + b_3 \times 4 + b_4 \times 5 = 34$.

- Ranges of variables:
  - $T \leq 5$
  - $0 \leq a_i \leq 10^9, \forall i$
  - $2 \leq K \leq N \leq 10^6$
- Subtask descriptions:
  - $N = 2$: 1 input file
  - $N \leq 2000$: 1 input file
  - $N \leq 10^5$, with the sequence $\{a_i | i = 0, 1, \ldots, N-1\}$ generated randomly: 1 input file
  - $1 \leq a_i \leq 10$: 1 input file
  - $K = 2$: 4 input files
  - $N \leq 10^6$: 4 input files
- Hints:
  - When K is equal to 2, then this problem is reduced to the problem of "running median" explained in the class.
  - To calculate the output, you can simply declare all the related variables as "unsigned integer" (unsigned int) in C++ and calculate the weighted sum directly without concerning overflow. In this manner, you do not need to perform the mod operation explicitly.

2. (30 pts) **Building hopping**: Given N buildings located in a row, with ith building having a height of h_i (starting from $i = 0$). If a person wants to move from building i to j, the following conditions must be satisfied:
   - $i < j$
   - $h_i < h_j$
   - $max(h_{i+1}, h_{i+2}, \ldots, h_{j-1}) \leq h_i$

   Now we have Q queries to be answered by your program. For the ith query, there are two persons in buildings a_i and b_i respectively, and they want to see each other is moving to the same building. So for each query, you need to output the building with the smallest index where two persons can see each other. If no such building exists, output -1 instead.

   More details:

   - Input format:
     - Line 1 contains an integer T which denotes the number of test cases in the file.
     - For each test case, the first line contains two integers N and Q, which are the numbers of buildings and queries, respectively.
     - In the next line, there are N space-separated integers $h_0, h_1, \ldots, h_{N-1}$, to indicate the height of the buildings.
     - In the next Q lines, the ith line contain two integer a_i and b_i indicating the building indices for two persons.
   - Output format:

        - For each query, output an integer as mentioned in the problem description.
    - Sample test file
        - Input file

```
2
6 4
1 2 1 3 4 2
0 1
0 2
1 3
3 5
2 1
1 4
0 0
```

        - Output file

```
1
3
3
-1
0
```

    - Ranges of variables:
        - $T \leq 7$
        - $1 \leq h_i \leq 10^9, \forall i$
        - $2 \leq N, Q$
        - $0 \leq a_i, b_i \leq N-1, \forall i$
    - Subtask descriptions:
        - $N \leq 2$, $Q \leq 4$: 1 input file
        - $N, Q \leq 5000$: 2 input file
        - $N \leq 3000$, $Q \leq 300000$: 2 input file
        - $N \leq 250000$, $Q \leq 500$: 2 input files
        - $N, Q \leq 250000$: 5 input files
    - Hints:
        - You may want to use the concept of lowest common ancestor if you want to solve all the subtasks.
        - If you use lowest common ancestor with binary search, be sure to set a smart upper bound to avoid TLE.

3. (40 pts) **Balanced parenthesis**: A balanced sequence of parentheses is good when each opening symbol has a corresponding closing symbol and the pairs of parentheses are properly nested. For example, "" , "()" , "()()", "((())()", "((()))" are good sequences, while ")", ")()", ")()))" are not.

   Now we receive a parenthesis sequence, but some of the symbols are missing and thus are denoted by "?". There are K "?" in total, and we need to replace "?" to either "(" or ")" to make the sequence good. However, the cost of replacing ith "?" with "(" is a_i, and with ")" is b_i (starting from $i = 0$). We want to minimize the cost when making the sequence good. If there is no way to make the sequence good, output -1 instead.

   More details:

    - Input format:
        - Line 1 is an integer T indicating the number of test cases in the file.
        - T test cases follow, where each test case is composed of 1 line of string denoting the sequence of "(", ")", and "?".
        - There are K lines follow, where K is the number of question mark ("?") in the sequence. The ith line contains two integers a_i and b_i,

which is the cost to replace the ith "?" with "(" and ")", respectively.

- Output format:
  - If there is no way to make the sequence good, print -1. Otherwise, print the minimum cost to make the sequence good.
- Sample test file
  - Input file

```
4
(
()(?
1 2
????
1 1
1 2
1 1
1 1
())?
3 4
```

  - Output file

```
-1
2
4
-1
```

- Ranges of variables:
  - $T \leq 20$
  - $1 \leq a_i, b_i \leq 10^9, \forall i$
  - $K \leq N$
- Subtask descriptions
  1. $K = 0, N \leq 10^6$: 2 input files
  2. $K = 1, N \leq 10^6$: 1 input file
  3. $K \leq 12, N \leq 1000$: 2 input files
  4. $K \leq 28, N = K, T \leq 20$: 2 input file
  5. $N \leq 2000, T \leq 15$: 1 input file
  6. $a_i = b_i = 1, N \leq 6 \times 10^5, T \leq 10$: 4 input files
  7. $N \leq 6 \times 10^5, T \leq 10$: 4 input files
- Hints:
  - You may want to use heap if you want to solve all the subtasks.

(Total score = 100)
