えびちゃん

**HCPC** 

November 21, 2019

### 今日の流れ

- 1. いつものコーナー
- 2. セグ木の導入
- 3. セグ木の実装
- 4. セグ木で解ける問題例
- 5. おわり

## 最近お勉強したもの (1/6)

えびちゃんが CHT の回以降に勉強したことたちです.興味があったら言ってね.

- simplex 法
  - 線形計画法を解くアルゴリズム
  - maximize  $c^{\top}x$  subject to  $Ax \leqslant b$  and  $x \geqslant 0$
  - 実用的には高速だが、最悪指数時間かかる
- wavelet matrix
  - 無限回言及したね
  - 整数の配列 α[n] に関するいろんなクエリに答えられる
  - i ∈ [s, t) のうちで a<sub>i</sub> ∈ [x, y) の個数は?
  - i ∈ [s, t) のうちで x が k 番目に現れる場所は?
  - i ∈ [s, t) のうちで k 番目に大きい値は?
  - i ∈ [s, t) のうちで x 以上の要素の総和は?
  - i ∈ [s, t) のうちの要素を頻出順に列挙して?

# 最近お勉強したもの(2/6)

- rolling hash  $\mathcal{O}$  hack
  - 弱いものなら誕生日攻撃で終了
  - tree attack というのがあるらしい
- Mo's algorithm
  - Mo Tao (中国人)
  - 区間クエリ処理に関する枠組み
  - 区間の左端の位置でバケットに分ける
- 並列二分探索
  - クエリ処理に関する枠組み(に限らないかも?)



# 最近お勉強したもの(3/6)

- 永続配列
  - バージョン管理できる配列みたいなの
- Karatsuba 法
  - 長さ  $\mathfrak{n}$  の畳み込みを  $O(\mathfrak{n}^{\log_2 3}) \subset O(\mathfrak{n}^{1.59})$  時間で行う
  - FFT と違って原始 n 乗根などがいらないので抽象化が楽 そう?
- Fibonacci heap
  - amortized O(1) で優先度を高くできる
  - Dijkstra 法や Prim 法で pop の回数を減らせて,
     O(|E| + |V| log |V|) にできる
  - 改良版もいろいろある

# 最近お勉強したもの (4/6)

- disjoint sparse table
  - 静的なモノイドの配列で、O(1) 時間で任意区間 fold できる
  - 空間は n ⋅ (log<sub>2</sub> n + O(1)) 要素ぶん
- sliding window aggregation
  - モノイドの両端キューで、ならし O(1) 時間で fold できる
  - 空間は 2n 要素ぶん
- word-level parallel
  - word あたり w bits あるので w bits をまとめて演算できる
  - あるいは  $w^{1/2}$  個の  $w^{1/2}$  bits の整数をまとめるとか
  - 最上位ビットを O(1) で求められたりする



# 最近お勉強したもの(5/6)

- WQS binary search (trick from Aliens)
  - Qingshi Wang (中国人)
  - 関数がいい性質を満たすときのテク
  - 回数に関する制約をペナルティで置き換える
- removable CHT
  - O((log n)³) でできるらしい
- x-fast trie
  - 整数 [0, U) に関する集合演算を O(log log U) 時間とかで
- 整数除算最適化
  - 除算は重くて加減乗算・ビット演算は軽い
  - $(n / 5) == ((n * 0xCCCCCCD_ju) >> 34)$
  - これは 32-bit 符号なし整数の例

# 最近お勉強したもの(6/6)

- monotone minima
  - monotone な m × n 行列について各行の最小値を O(n log m) 時間で求める
- SMAWK
  - totally monotone な m × n 行列について各行の最小値を O(m+n) 時間で求める



### 非再帰セグ木の導入

ここから非再帰セグ木の話.

再帰セグ木を知っている人へ:

再帰セグ木を stack などで無理やり書いたものではないです.

たぶんボトムアップセグ木とか呼ぶ方が適切?



配列  $a=(a_0, a_1, \ldots, a_{n-1})$  が与えられる. 次のクエリを q 回処理してね.  $1 \le n \le 10^5, 1 \le q \le 10^5$ .

#### クエリセット(かんたん)

- i が与えられて、αi の値を答える
- (i, x) が与えられて、a<sub>i</sub> を x に書き換える

配列  $a = (a_0, a_1, \ldots, a_{n-1})$  が与えられる. 次のクエリを q 回処理してね.  $1 \le n \le 10^5, 1 \le q \le 10^5$ .

#### クエリセット(かんたん)

- i が与えられて、α<sub>i</sub> の値を答える
- (i, x) が与えられて、a<sub>i</sub> を x に書き換える

配列を知っていますか?

配列  $\mathfrak{a}=(\mathfrak{a}_0,\,\mathfrak{a}_1,\,\ldots,\,\mathfrak{a}_{\mathfrak{n}-1})$  が与えられる. 次のクエリを  $\mathfrak{q}$  回処理してね.  $1\leqslant\mathfrak{n}\leqslant 10^5,\,1\leqslant\mathfrak{q}\leqslant 10^5$ .

#### クエリセット(むずかしい)

- ullet (l, r) が与えられて,  $\sum_{i=1}^{r-1} a_i$  の値を答える
- (i, x) が与えられて、αi を x に書き換える



配列  $a=(a_0, a_1, \ldots, a_{n-1})$  が与えられる. 次のクエリを q 回処理してね.  $1 \le n \le 10^5, 1 \le q \le 10^5$ .

#### クエリセット(むずかしい)

- ullet (l, r) が与えられて,  $\sum_{i=1}^{r-1} a_i$  の値を答える
- (i, x) が与えられて、αi を x に書き換える

for 文を知っていますか?

配列  $a=(a_0, a_1, \ldots, a_{n-1})$  が与えられる. 次のクエリを q 回処理してね.  $1 \le n \le 10^5, 1 \le q \le 10^5$ .

#### クエリセット(むずかしい)

- (l, r) が与えられて、 $\sum_{i=1}^{r-1} a_i$  の値を答える
- (i, x) が与えられて、ai を x に書き換える

for 文を知っていますか?

↑計算量を知っていますか?

配列  $\mathfrak{a}=(\mathfrak{a}_0,\,\mathfrak{a}_1,\,\ldots,\,\mathfrak{a}_{\mathfrak{n}-1})$  が与えられる. 次のクエリを  $\mathfrak{q}$  回処理してね.  $1\leqslant\mathfrak{n}\leqslant 10^5,\,1\leqslant\mathfrak{q}\leqslant 10^5$ .

#### クエリセット(むずかしい)

- (l, r) が与えられて, $\sum_{i=1}^{r-1} a_i$  の値を答える
- (i, x) が与えられて、ai を x に書き換える

for 文を知っていますか?

- ↑計算量を知っていますか?
- →セグ木を使うと  $O(n + q \log n)$  で解ける.

### 基本アイディア

隣同士の和を求める. そのペアごとに, 和を求める. というのを繰り返す.

a	$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$									
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$										
a <sub>0</sub> -	$a_0 + a_1$ $a_2 + a_3$			a <sub>4</sub> -	⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>			
$a_0$ $a_1$ $a_2$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$										

Figure: セグ木の概念図

### 基本アイディア

隣同士の和を求める. そのペアごとに, 和を求める. というのを繰り返す.

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$								
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$								
a <sub>0</sub> -	$a_0 + a_1$ $a_2 + a_3$			$\mathfrak{a}_4$ -	⊢ α <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>	
$a_0$ $a_1$ $a_2$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$								

Figure:  $a_0 + a_1 + \cdots + a_6$  に対応する区間

隣同士の和を求める. そのペアごとに, 和を求める. というのを繰り返す.

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$										
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$										
a <sub>0</sub> -	⊦ a <sub>1</sub>	$\mathfrak{a}_2$	$+a_3$	$a_4 +$		a <sub>6</sub> -	⊢ a <sub>7</sub>			
ao	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	<b>a</b> 4	<b>a</b> 5	$\mathfrak{a}_6$	<b>a</b> 7			

Figure: a<sub>2</sub> が関与する区間

### 木の表現

配列を用いる. 根の添字を1として, 幅優先順に番号をつける.

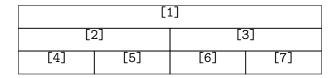


Figure: セグ木の添字

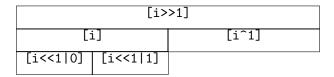


Figure: 親子の添字(綺麗!)

### 木の表現(悪い例)

配列を用いる. 根の添字を0として, 幅優先順に番号をつける.

[0]						
[1	[]	[2]				
[3]	[4]	[5]	[6]			

Figure: セグ木の添字

親子の添字を求めるのが面倒なことがわかる(はず). なんでも 0-indexed にしたがる人は考え直した方がよい.

- 少ない区間で所望の区間をカバーしたい。
- 上の要素ほど大きな区間をカバーする.
- → できるだけ上の要素で足したい.

ある要素を足さなきゃいけない条件は?

→ その親の要素では所望の区間をはみ出してしまうとき.



所望の区間を半開区間で表す.



Figure: 半開区間 [i, j) を意味する矢印

区間の左端・右端と、左の子・右の子の4パターンを考える.



まず左端について考える.



Figure: 左端・左の子

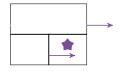


Figure: 左端・右の子



👚 は,その要素を足すことを意味する.足したら区間をずらす.

次に右端について考える.



Figure: 右端・左の子



Figure: 右端・右の子



👚 は,その要素を足すことを意味する.足したら区間をずらす.

要素数 n の配列に対応するセグ木で、[l, r) の和を求める.

```
x = 0;
l += n, r += n;
while (l < r) {
   if (l & 1) x += c[l++];
   if (r & 1) x += c[--r];
   l >>= 1, r >>= 1;
}
return x;
```

交換法則を仮定したくないときも,少しの手直しで対処可能(読者への課題).

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$										
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$										
a <sub>0</sub> -	+ a <sub>1</sub>	a <sub>2</sub> -	- a <sub>3</sub>	a <sub>4</sub> -	⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ α <sub>7</sub>			
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$										

Figure: [1, 7) の和を求めてみる

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$										
ao	$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$									
a <sub>0</sub> -	$a_0 + a_1$ $a_2 + a_3$				$a_4 + a_5$ $a_6 + a_7$					
$a_0$ $a_1$ $a_2$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$										

Figure: [1,7) の和を求めてみる

Figure: [1, 7) の和を求めてみる

### 変更するクエリ

自分の項が関係するのは祖先ノードのみ.

→親を辿っていけばよい.

```
i += n;
c[i] += x;
while (i > 1) {
   i >>= 1;
   c[i] = c[i<<1|0] + c[i<<1|1];
}</pre>
```

非再帰セグ木

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$									
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$									
a <sub>0</sub> -	$a_0 + a_1$ $a_2 + a_3$				⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>		
$a_0$ $a_1$ $a_2$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$									

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$									
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$									
a <sub>0</sub> -	$a_0 + a_1$ $a_2 + a_3$				⊦ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>		
$a_0$ $a_1$ $x$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$									

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$								
$a_0 + a_1 + a_2 + a_3$ $a_4 + a_5 + a_6 + a_7$								
a <sub>0</sub> -	$a_0 + a_1$ $x + a_3$				⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>	
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$								

$a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$								
$a_0 + a_1 + x + a_3$ $a_4 + a_5 + a_6 + a_7$								
a <sub>0</sub> -	$a_0 + a_1$ $x + a_3$			a <sub>4</sub> -	⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>	
$a_0$ $a_1$ $x$ $a_3$ $a_4$ $a_5$ $a_6$ $a_7$								

$a_0 + a_1 + x + a_3 + a_4 + a_5 + a_6 + a_7$										
$a_0 + a_1 + x + a_3$ $a_4 + a_5 + a_6 + a_7$										
a <sub>0</sub> -	⊦ a <sub>1</sub>	x +	· a <sub>3</sub>	a <sub>4</sub> -	⊢ a <sub>5</sub>	a <sub>6</sub> -	⊢ a <sub>7</sub>			
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$										

# トのコードは要素数 n を二べキにしなくても動く.

O(log n) 個の完全二分木のセグ木が存在していると見なせる.

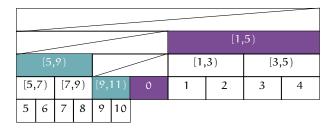


Figure: 11 要素の配列に対応するセグ木

### 驚愕の事実?

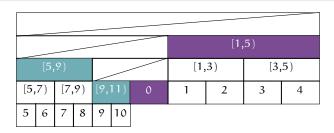


Figure: 配列に詰め込んだ形

	[1,5)				[5,9)					
	[1,	3)	[3,5)		[5,7)		[7,9)		[9,11)	
0	1	2	3	4	5	6	7	8	9	10

Figure: 実質的に扱う概念

#### 抽象化

和に限ったデータ構造ではなく,モノイドならなんでもよい.

- $\bullet (a \circ b) \circ c = a \circ (b \circ c)$
- $\exists e \text{ s.t. } a \circ e = e \circ a = a$

 $(\{i, i+1, \ldots, i+n\}, min)$  はモノイドを成す.単位元はi+n.

中級者向け:ロリハは連接についてモノイドを成す.

基数 b, 法 p と する.

長さ  $n_1$ ,  $n_2$  の文字列のハッシュ値  $h_1$ ,  $h_2$  について,

$$(n_1, h_1) \circ (n_2, h_2) = (n_1 + n_2, h_1 \cdot b^{n_2} + h_2).$$

非再帰セグ木

## 抽象化

モノイドになるようにうまく言い換えられるとうれしい.

たとえば、正しい括弧列を考えてみる.

開き括弧を +1, 閉じ括弧を -1 として, 正しい括弧列である条件をうまく表せないか考えてみよう.



# セグ木上のにぶたん (セグメント木上の二分探索)

非負整数の配列  $a = (a_0, a_1, \ldots, a_{n-1})$  について,先頭からの 累積和がxを超えない限界はどこ? 形式的には、以下のものが知 りたくなってみる.

max. s s.t. 
$$\sum_{i=0}^{s-1} a_i \leqslant x.$$

 $s \in [0, n+1)$  を決め打ちして, [0, s) での和を求めてにぶたん? →これをすると  $O((\log n)^2)$ .

#### セグ木上のにぶたん

セグ木をよく見ると、すでに半々に分けられた構造をしているこ とがわかる. →左右どちらに辿るかを判定すればよい.





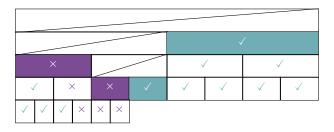
Figure: 左の子を辿る

Figure: 右の子を辿る

左の子を足しても条件を満たすなら、足して右の子を辿る、そう でなければ、そのまま左の子を辿る.

 $\checkmark$  はそこまで足しても条件を満たすことを $, \times$  はそこまで足すと 条件を満たさないことを意味する.

完全二分木でないときは次のような感じになっている.



000000000000

Figure: 11 要素の配列に対応するセグ木

木は  $O(\log n)$  個しかないので、どの根から始めるかを線形探索できる、任意位置を始点とすることもできる(考えてみよう).

- 辺同ン素要も另コ专ろる次を3kの(+,1] ,制財专べる見コ专ろをかみずぶコで(+,1] 間図:イビコ

# 動的配列の模倣

以下のクエリを処理してね.

#### クエリセット(むずかしそう)

- 集合に要素 x を追加する
- 集合から要素 x を削除する
- 小さい方から k 番目の要素を答える

ただし、 $0 \le x < 10^5$ . x を先読み可能なら  $|x| \le 10^9$  でも.

0000000000000

## 動的配列の模倣

以下のクエリを処理してね.

#### クエリセット(むずかしそう)

- 集合に要素 x を追加する
- 集合から要素 x を削除する
- 小さい方から k 番目の要素を答える

ただし,  $0 \le x < 10^5$ . x を先読み可能なら  $|x| \le 10^9$  でも.

std::set ではk番目へのアクセスはできない.

0000000000000

## 動的配列の模倣

以下のクエリを処理してね.

#### クエリセット(むずかしそう)

- 集合に要素 x を追加する
- 集合から要素 x を削除する
- 小さい方から k 番目の要素を答える

ただし、 $0 \le x < 10^5$ . x を先読み可能なら  $|x| \le 10^9$  でも.

std::set ではk番目へのアクセスはできない.

AVL 木や赤黒木を貼るしかない...? →セグ木でできます.

つよい bit set で 64 倍高速化してもいいです.

 $(\mathbb{N}, +)$  を管理するセグ木を使う.

要素 i が集合に入っているとき  $a_i = 1$ , otherwise  $a_i = 0$  とする.

追加・削除は簡単に行える(0または1を代入すればよいので).

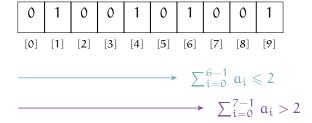
k 番目の要素は、このセグ木で和が k になる添字に対応するので、 セグ木上でのにぶたん。

図があった方がわかりやすそう. 次ページ.



## 動的配列の模倣

集合  $\{1, 4, 6, 9\}$  に対応する配列(実際にはセグ木で管理する). ここから 0-indexed で 2 番目の要素を探してみる.



(セグ木上のにぶたんができて)6が答えだとわかる.

転倒数というのを求めてみよう.

#### 転倒数

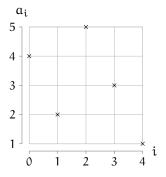
配列  $a = (a_0, a_1, \ldots, a_{n-1})$  について、次の条件を満たす (i, j)のペアの個数を転倒数と呼ぶ.

- i < j
- $a_i > a_i$

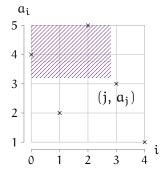
ウェーブレット行列はこのクエリに直接答えられるが...

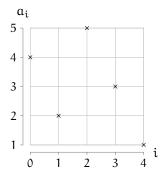
→セグ木でできます.

例として、配列  $\alpha = (4, 2, 5, 3, 1)$  は次のように見なせる.

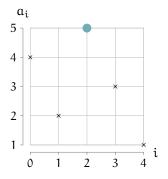


点  $(j, a_i)$  に関して、i < j かつ  $a_i > a_i$  は次の領域に相当する.

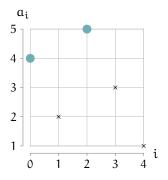




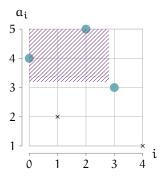
$$b = (0, 0, 0, 0, 0), x = 0$$



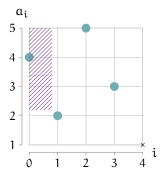
$$b = (0, 0, 1, 0, 0), x = 0 + 0$$



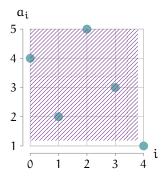
$$b = (1, 0, 1, 0, 0), x = 0 + 0 + 0$$



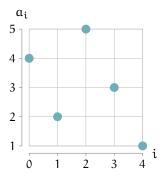
$$b = (1, 0, 1, 1, 0), x = 0 + 0 + 0 + 2$$



$$b = (1, 1, 1, 1, 0), x = 0 + 0 + 0 + 2 + 1$$



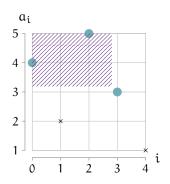
$$b = (1, 1, 1, 1, 1), x = 0 + 0 + 0 + 2 + 1 + 4$$



$$b = (1, 1, 1, 1, 1), x = 7$$

以上により, (4, 2, 5, 3, 1) の転倒数は7だとわかった.

 $a_i$  の順で行うことで、 $a_i$  の大小関係と処理順が対応づけられる のがうれしい.



実装の際は, i > j かつ  $a_i < a_j$  とみて,  $a_i$  の小さい方から走査し て、右下の点の個数を数えてもよい、

同じ値の組が入っているときに、それらを余分に数えないように 注意しよう. たとえば, (x, x) の転倒数が正しく 0 となるか確認 しよう.

少し考えると, 同じような方法で最大増加部分列 (LIS) も平面走 査で解けることがわかる.これは、蟻本 pp. 63-65 に載っている 方法とは異なる. ∞ に相当する要素を用意しなくていいのが利点 としてありそう. 計算量は  $\Theta(n \log n)$  のまま.

## 平面走査(応用例)

たとえば,ある条件を満たす区間 [l, r) の個数を求めたいとする. ある f が存在して、その条件を満たすことと  $f(l) \leq f(r)$  が同値で あるとする.

→これは平面走査で解ける形をしているので、平面走査で解ける.

#### 問題例

条件 P と配列  $a = (a_0, a_1, \ldots, a_{n-1})$  が与えられる. 次の条件 を満たす区間 [1, r] の個数を数えよ.ただし  $1 \le n \le 10^5$ .

Pを満たす要素の方が満たさない要素より真に多い Ρ(α;) の計算は高速にできると仮定してよい.

#### まとめ

セグ木では、モノイドの配列への更新と fold を高速に行える. モノイドを利用できる形に言い換えられるとうれしい。 更新順序をうまく変えて平面走査するテクも有効.

#### 非再帰セグ木の実装:

- 内部実装で 0-indexed にこだわるべきでない
- 二ベキに丸める必要もない

ところで再帰セグ木の紹介をしていませんね。

## 問題たち

- https://onlinejudge.u-aizu.ac.jp/courses/library/3/DSL/al1/DSL\_2\_A
- https://yukicoder.me/problems/no/875
- https://atcoder.jp/contests/abc127/tasks/abc127\_f
- https: //atcoder.jp/contests/bitflyer2018-final/tasks/bitflyer2018\_final\_c
- https://atcoder.jp/contests/arc033/tasks/arc033\_3
- https://yukicoder.me/problems/no/877
- https://atcoder.jp/contests/past202004-open/tasks/past202004\_n
- https://atcoder.jp/contests/chokudai\_S001/tasks/chokudai\_S001\_j
- https://atcoder.jp/contests/arc075/tasks/arc075\_c
- https://atcoder.jp/contests/arc101/tasks/arc101 b
- https://atcoder.jp/contests/dp/tasks/dp\_q

#### 発展的な話

静的なら disjoint sparse table で O(1) 回の演算で fold 可能.

よい性質を満たすとき,fold に加えて区間の更新も O(log n) 回の 演算で可能、遅延伝播セグメント木(遅延セグ木).

- 区間 min・区間加算は可能
- 区間和・区間加算は可能
- 区間和・区間 min 更新は不可能 → segment tree beats!

beats は謎、なんかいろいろできるらしい.

