1 Static segment tree

Let $a_0 < a_1 < a_2 < \cdots < a_n$ be a sequence of real numbers, let $A = \{a_0, \ldots, a_n\}$. Consider a binary tree T where the leaves correspond to "atomic" intervals $[a_i, a_{i+1}]$ for $i \in \{0, \ldots, n-1\}$ and internal nodes correspond to unions of primitive intervals. If a node v has children u, w with corresponding intervals $[\ell_u, r_u]$ and $[\ell_w, r_w]$ then $r_u = \ell_w$ and the interval corresponding to v is $[\ell_v, r_v]$, where $\ell_v = \ell_u$ and $r_v = r_w$

Definition 1.1 Given an interval I with endpoints in A its **canonical decomposition** is a disjoint union of intervals corresponding to the nodes of the tree where the number of intervals is minimized.

The canonical decomposition is unique and can be computed as follows.

`canonical`$(v, [x, y])$ invariant is $\ell_v \leq x < y \leq r_v$
  `if` $[x, y] = [\ell_v, r_v]$ `then return` $[\ell_v, r_v]$
  let u, w be the children of v; let $m = y_u$ (note $m = x_w$)
  `if` $x < m < y$ `then return canonical`$(u, [x, m])$, `canonical`$(w, [m, y])$;
  `if` $y \leq m$ `then return canonical`$(u, [x, y])$;
  `if` $m \leq x$ `then return canonical`$(w, [x, y])$;

Lemma 1.2 *The number of calls to* `canonical` *is at most* 4 *times the depth of the tree.*

Proof :
We will assign a value to call of `canonical`$(v, [x, y])$ as follows. If $\ell_v = x$ and $r_v = y$ then the value is 2. If either $\ell_v = x$ or $r_v = y$ then the value is 1. Otherwise the value is 0. We will say that a call **branches** if we are in the case $x < m < y$. Note that if a call does not branch then it recursively makes at most one call with at least as large value. Note that

- A value 2 call does not produce any further calls.
- If a call branches then the two calls each have value at least 1.
- If a call of value 1 branches then one of the two calls has value 2 and the other one has value at least 1.

From this, by induction, it follows that on a level of the tree we can have

- one call of value 0, or
- at most two calls of value 1 and at most two calls of value 2.

Hence the total number of calls is at most four times the depth of the tree. ∎

We will maintain the following additional information at each node of the tree. Suppose that intervals $I_1, \ldots, I_k$ have been added to the segment tree.

- `count`(v): is the number of $i \in [k]$ such that the canonical decomposition of I_i contains $[\ell_v, r_v]$ (that is, the interval corresponding to v).
- `measure`(v): is the measure (total length) of the union of canonical intervals for $I_1, \ldots, I_k$ that are contained in the subtree rooted by v.

Lemma 1.3 *We have the following properties.*

- *If* $\texttt{count}(v) > 0$ *then* $\texttt{measure}(v) = r_v - \ell_v$.
- *If* $\texttt{count}(v) = 0$ *then* $\texttt{measure}(v) = \texttt{measure}(u) + \texttt{measure}(w)$, *where* u *and* w *are the children of* v.
- *For the root* r *of the tree* $\texttt{measure}(r)$ *is the measure (total length) of the union of intervals.*

The following allows to add/remove intervals from the collection (for add $c = +1$, for remove $c = -1$).

```
UPDATE(v, [x, y], c)            invariant is ℓ_v ≤ x < y ≤ r_v
     if [x, y] = [ℓ_v, r_v] then
         count(v)+ = c
    else
         let u, w be the children of v; let m = y_u (note m = x_w)
         if x < m < y then UPDATE(u, [x, m], c); UPDATE(w, [m, y], c);
         if y ≤ m then UPDATE(u, [x, y], c);
         if m ≤ x then UPDATE(w, [x, y], c);

     if count(v) = 0 then
         if v is a leaf then measure(v)=0 else measure(v)=measure(u)+measure(w)
    else measure(v)=r_v − ℓ_v
```