Разреженная таблица

ceagest

1 Разреженная таблица (Sparse Table)

Определение 1.1. *Пусть дан массив $\mathcal{A}$. Разреженная таблица или Sparse Table — двумерная структура данных $ST[i][j]$, построенная на бинарной операции $\mathcal{F}$, для которой выполнено следующее:*

$$ST[i][j] = \mathcal{F}(\mathcal{A}[i], \mathcal{A}[i+1], ..., \mathcal{A}[i+2^j-1]), \quad j \in \{1, ..., \log_2 n\}$$

Замечание 1.1. *Так как в определении 1.1 операция $\mathcal{F}$ является бинарной, то под записью $\mathcal{F}(\mathcal{A}[i], \mathcal{A}[i+1], ..., \mathcal{A}[i+2^j-1])$ подразумевается следующее: $\mathcal{F}(\mathcal{A}[i], \mathcal{F}(\mathcal{A}[i+1], \mathcal{F}(..., \mathcal{F}(\mathcal{A}[i+2^j-2], \mathcal{A}[i+2^j-1])...))$.*

1.1 Построение Sparse Table

Простой метод построения таблицы заключён в следующем рекуррентном соотношении:

$$ST[i][j] = \begin{cases} \mathcal{F}(ST[i][j-1], ST[i+2^{j-1}][j-1]), & \text{если } j > 0 \\ \mathcal{A}[i], & \text{если } j = 0 \end{cases}$$

Замечание 1.2. *Для лучшей работы кешей лучше хранить матрицу, у которой мало длинных строк (тогда для ответа на запрос будут браться значения из одного массива).*

1.2 Ответ на запрос

Заметим, что в этой таблице хранятся результаты функции $\mathcal{F}$ на всех отрезках, длины которых равны степеням двойки. Выполним сначала предподсчет, суть которого в вычислении массива E такого, что $E[j] = \lfloor \log_2 j \rfloor$. Теперь заметим, что для отрезка $[l, r]$ верно, что:

$$\mathcal{F}(\mathcal{A}[l], \mathcal{A}[l+1], ..., \mathcal{A}[r]) = \mathcal{F}(ST[l][j], ST[r-2^j+1][j]), \quad j = E[r-l+1]$$

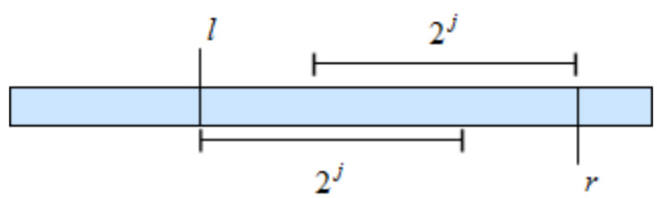

Рис. 1: Отрезки, которые мы берём для получения ответа на отрезке $[l, r]$

1.3 Ресурсы

- В таблице хранятся результаты функции $\mathcal{F}$ на всех отрезках, длины которых равны степеням двойки. Однако $\forall j \in \{0, ..., \log_2 n\}$ таких отрезков не более n, откуда потребляемая память составит $\mathcal{O}(n \log n)$.
- Теперь время построения. Заметим, что каждая ячейка пересчитывается за $\mathcal{O}(1)$, откуда время построения $\mathcal{O}(n \log n)$.
- И последнее – ответ на запрос. Заметим, что это всего лишь вычисление функции от двух значений, что работает за $\mathcal{O}(1)$.

1.4 Требования к $\mathcal{F}$

Бинарная операция $\mathcal{F}$ должна удовлетворять следующим условиям:

1. Идемпотентность
2. Ассоциативность
3. Коммутативность

Определение 1.2. *Пусть* $\mathbb{X}$ *– абстрактное множество,* $\mathcal{B}\colon \mathbb{X} \times \mathbb{X} \to \mathbb{X}$. *Будем говорить, что* $\mathcal{B}$ *– идемпотентна, если* $\forall x \in \mathbb{X} \quad \hookrightarrow \quad \mathcal{B}(x, x) = x$.