

Подозрительные отчеты

Автор задачи и разработчик: Мария Жогова

Рассмотрим классическую динамику для задачи нахождения *наибольшей общей последовательности*. В данном случае мы хотим проверить, что сокращенный отчет целиком «входит» в полный отчет с указанными дополнительными ограничениями. Будем хранить $\text{dp}[i][j]$ — могут ли первые i столбцов гистограммы полного отчета содержать первые j столбцов противозаконного отчета, чтобы i -й соответствовал j -му.

При пересчете такой динамики достаточно перебрать k — предыдущий столбец первой гистограммы, соответствующий $j - 1$ -му столбцу второй. На него накладываются следующие ограничения: $s_k - t_{j-1} = s_i - t_j$ (то, что гистограмма была обрезана по горизонтальной линии, означает, что разности высот столбцов одинаковы) и $s_x \leq s_i - t_j$ для всех $k < x < i$ (так как все элементы между выбранными столбцами должны быть не выше линии разреза). Таким образом,

$$\text{dp}[i][j] = \bigvee_{\substack{k < i \\ s_k - t_{j-1} = s_i - t_j \\ \max_{k < x < i} s_x < s_i - t_j}} \text{dp}[k][j-1]$$

Такой пересчет в наивной реализации работает за $\mathcal{O}(nm^2)$. Попробуем его оптимизировать.

Во-первых, несложно заметить, что нам надо ограничиться только такими k , для которых $s_k = s_i - t_j + t_{j-1}$. А также можно заметить, что достаточно перебрать не все такие k , а только максимальный подходящий. Действительно, чем меньше k , тем больше максимум на отрезке s_{k+1}, \dots, s_{i-1} , а значит если через какой-то меньший k можно обновиться, то и для максимального все условия будут выполнены. А также, любая НОП, заканчивающаяся в меньших k , может быть модифицирована перемещением последнего столбца в больший k без потери корректности. Поэтому достаточно пересчитать $\text{dp}[i][j] = \text{dp}[k][j-1]$ для максимального $k < i$, для которого $s_k - t_{j-1} = s_i - t_j$ и $\max_{k < x < i} s_x < s_i - t_j$.

Последнее условие либо выполнено для максимального подходящего по разности столбцов k , либо не выполнено вообще ни для какого. Найти максимальный подходящий k можно, запомнив для каждого числа в массиве последнее его вхождение. В данном случае мы знаем, что мы ищем $s_k = s_i - t_j + t_{j-1}$, значит достаточно взять $k = \text{last}[s_i - t_j + t_{j-1}]$. Это все можно поддерживать с помощью, например, хэш-таблиц, пока идем слева-направо.

Осталось только для полученного k проверять условие на максимум между ним и i . Искать максимум на интервале (k, i) можно с помощью *разреженной таблицы*. Предподсчитав ее на гистограмме s , можно проверять, ограниченность максимума на отрезке за $\mathcal{O}(1)$. Итоговая сложность: $\mathcal{O}(n \log n + nm)$