Отрезки

1 Постановка задачи

Дана система из n отрезков на числовой прямой:

$$\{[l_i,\, r_i] \mid i = 1, 2, \ldots, n\},$$

где $l_i \le r_i$. Требуется выбрать наименьшее по мощности множество точек $S \subset \mathbb{R}$ так, чтобы каждый отрезок $[l_i, r_i]$ содержал хотя бы одну точку из S.

2 Описание алгоритма

Для решения задачи удобно воспользоваться жадной стратегией. Сначала мы сортируем все отрезки по неубыванию их правых концов. После этой сортировки мы просматриваем отрезки в порядке увеличения значения r_i. При обработке каждого отрезка проверяем, содержится ли в нём уже ранее выбранная точка; если нет, добавляем в результат именно правый конец текущего отрезка и тем самым помечаем его как «покрытый». В дальнейшем все отрезки, которые содержат эту новую точку, будут автоматически считаться закрытыми, и мы их просто пропустим.

3 Обоснование корректности

Жадный выбор правого конца самого «раннего» (т.е. с наименьшим правым концом) непокрытого отрезка гарантирует оптимальность решения. Действительно, если какой-либо отрезок имеет минимальный правый конец r, то любая точка, лежащая внутри этого отрезка, не сможет «закрыть» какие-то отрезки с меньшим правым концом (поскольку их нет), но замена любой другой точки на r при покрытии этого отрезка не ухудшит покрытия остальных: все отрезки, содержащие исходную точку, будут по-прежнему содержать точку r, так как r лежит правее. Следовательно, выбор r не хуже любого другого варианта, и, удалив все отрезки, покрытые точкой r, мы получаем задачу того же типа для оставшихся отрезков. Повторяя это рассуждение, получаем, что жадная процедура захватывания правых концов до тех пор, пока остаются непокрытые отрезки, даст число точек, не превосходящее число точек в любом оптимальном покрытии.

4 Асимптотическая оценка

После сортировки n отрезков по правым концам, которая занимает $O(n \log n)$, мы проходим по полученному списку за $O(n)$. Итоговая временная сложность алгоритма равна

$$O(n \log n) + O(n) = O(n \log n).$$