

## range-tree

- محدودیت زمان: ۲ ثانیه
- محدودیت حافظه: ۲۵۶ مگابایت

در پروژه پایانی، ما قصد پیاده سازی یکی از ساختمان داده های مکانی (spatial) برای  $\mathbb{R}^2$  را داریم. نام این ساختمان داده range-tree است. این درخت یکی از پرکاربردترین ساختمان داده ها در بانک های اطلاعاتی مکانی است. این ساختمان داده برای هر بعدی از فضا خوش تعریف است اما برای سادگی پیاده سازی فرض میکنیم فضا به یک صفحه اقلیدسی محدود شده است.

قبل از اینکه range-tree در دو بعد را توضیح دهیم ابتدا در یک بعد را توضیح میدهیم و بعدا بیان خواهیم کرد چگونه می توان آن را به دو بعد (و حتی ابعاد بالاتر) گسترش داد. فرض کنید  $n$  نقطه در صفحه داده شده است. ابتدا فقط مقادیر  $x$  این نقاط را (که مجموعه ای از اعداد بر روی محور  $x$  می باشد) در نظر می گیریم. ما قصد داریم ساختمان داده ای بسازیم که آماده جواب دادن به این سوال باشد: برای یک بازه داده شده مانند  $I = [a, b]$  بر روی محور  $x$ ، مقدار  $x$  کدام یک از نقاط اولیه در بازه  $I$  قرار دارد. درخت اولیه برای یک بعد (محور  $x$ ) به این صورت ساخته می شود. ابتدا همه نقاط را بر اساس مقدار  $x$  مرتب میکنیم. سپس دو تا دو تا، به ترتیب به یک گره پدر وصل کرده و مقدار گره پدر، برابر خواهد بود با مقدار بزرگترین  $x$  در زیردرخت چپ آن. اگر در ابتدا تعداد فرد بود، همین عملیات دودویی را برای گره آخر با گره پدر جفت آخر تکرار میکنیم. همین عملیات را تکرار میکنیم تا به یک گره برسیم. نتیجه درخت بالانسی خواهد بود (مانند AVL) که در زمان  $O(\log n)$  می توان در آن جای یک مقدار  $x$  را (همانند BST) پیدا کرد. به این درخت، درخت rang-tree بر روی محور  $x$  میگوییم و با  $R^x$  نشان می دهیم.

بعد از ساخت درخت برای محور  $x$  به سراغ محور دوم می رویم. برای هر گره میانی مانند  $v$  در درخت  $R^x$  (گره ای که برگ نباشد)، یک اشاره گر وجود دارد به درخت دومی. این درخت دوم برابر است با درخت range-tree از تمام برگ هایی که در زیردرخت  $v$  قرار داشتند، اما بر اساس مقادیر  $y$ . به این درخت، درخت rang-tree بر روی محور  $y$  برای گره  $v$  میگوییم و با  $R_v^y$  نشان می دهیم.

یک مثال در شکل زیر نشان داده شده است. در این مثال ۷ نقطه وجود دارد:

$(1, 1), (9, 3), (2, 7), (7, 2), (3, 5), (5, 8), (6, 6)$

برای انجام جستجو در این درخت برای فضای دو بعدی (که در زمان  $O(\log^2 n + k)$  انجام می شود و  $k$  تعداد نقاط داخل جواب نهایی هستند) یک مستطیل مانند  $R = [(x_1, y_1), (x_2, y_2)]$  داده شده است و ما قصد داریم تمام نقاط داخل یا روی این مستطیل را گزارش کنیم (نقطه اول گوشه چپ پایین و نقطه دوم گوشه راست بالا است). برای جواب دادن به این سوال ابتدا روی  $R^x$  شروع به جستجو می کنیم. گره ای که مقدارش از  $x_1$  بزرگتر باشد و از  $x_2$  کوچکتر باشد را

گره  $u$  می نامیم (اگر در هنگام جستجو مقدار گره میانی با مقدار  $x_2$  برابر شد برای ادامه جستجو به سمت چپ می رویم). تمام گره هایی که بین دو مسیر جستجو برای  $x_1$  و  $x_2$ ، از ریشه تا برگ، قرار میگیرند دارای اهمیت هستند و تمام نقاط در زیردرخت ها امکان حضور در جواب را دارند و باید مقدار  $y$  آنها مورد بررسی قرار بگیرد. پس به سراغ هر کدام از این درخت ها مانند  $R_u^y$  می رویم. در این درخت ها شروع به جستجو کرده و گره ای را که مقدارش از  $y_1$  بزرگتر باشد و از  $y_2$  کوچکتر باشد را گره  $t$  می نامیم. تمام برگ هایی که بین دو مسیر جستجو برای  $y_1$  و  $y_2$ ، از ریشه تا برگ، قرار می گیرند، نقاطی هستند که در جواب نهایی وجود دارند. در انتها فقط یک مرتب سازی بر اساس مقادیر  $y$  نیاز هست تا با تست های کویرا منطبق شود (وگرنه ضرورتی برای مرتب بودن نتیجه وجود نداشت).

برای مطالعه بیشتر می توانید منابع زیر را مطالعه کنید، اما دقت کنید کد کپی نکنید و همچنین به دنبال ساختمان داده های دیگر نروید. در ضمن توصیه می شود از زبان های مناسب این کار مثل Cpp بهره ببرید (مسئولیت استفاده از زبان های دیگر با دانشجو است):

- <https://www.youtube.com/watch?v=xVka6z1hu-I>
- <http://www.cs.umd.edu/~meesh/cmsc420/ContentBook/FormalNotes/MountNotes/lecture20-rangetrees.pdf>
- <http://www.cs.uu.nl/docs/vakken/ga/slides5b.pdf>

راستی ویکیپدیا در مورد پیچیدگی زمانی اشکالاتی دارد (دقت کنید). زمان ساخت شما باید بیشتر از  $O(n \log n)$  زمان و حافظه مصرف نکند و جواب هر سوال را در  $O(\log^2 n + k)$  بدهد (زمان مرتب سازی آخر در نظر گرفته نشده است).

## ورودی

خط اول تعداد نقاط ورودی را نشان می دهد. در خط دوم مقادیر  $x$  نقاط ورودی و در خط سوم مقادیر  $y$  متناظر با مقادیر خط قبلی داده شده است. پس اولین ایکس و اولین وای با یکدیگر اولین نقطه را نمایش می دهد. خط چهارم تعداد عملیات های بعدی را مشخص می کند. از خط ۵ به بعد هر خط چهار مقدار را مشخص می کند که به ترتیب مختصات  $x_1, y_1, x_2, y_2$  برای یک مستطیل است. فرض کنید تعداد ورودی و تعداد عملیات ها از  $10^4$  بیشتر نخواهد بود.

## خروجی

برای هر عملیات از خط ۵ به بعد در یک خط مقادیر  $x$  و در خط بعدی مقادیر  $y$  نقاطی که در داخل مستطیل می افتند را چاپ کنید. دقت کنید نقاط خروجی به ترتیب مقدار  $y$  از کوچک به بزرگ مرتب شده باشند. اگر جواب خالی بود None را چاپ کند.

## مثال

### ورودی نمونه ۱

```

3
1 2 3
1.2 2.1 3.3
2
0 0 2.1 1.4
1.5 2.2 2.8 5
    
```

خروجی نمونه ۱

```

1
1.2
None
    
```

ورودی نمونه ۲

```

7
1 9 2 7 3 5 6
1 3 7 2 5 8 6
3
0 0 5 5
8 1 10 5
0 0 10 10
    
```

خروجی نمونه ۲

```

1 3
1 5
9
3
1 7 9 3 6 2 5
1 2 3 5 6 7 8
    
```