

کوئیز طراحی الگوریتم ها

1. مرتب سازی مبنایی (Radix) را تعریف نموده و دنباله زیر را با نمایش مراحل بنویسید .

| | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 324 | 239 | 879 | 358 | 225 | 137 | 317 | 416 | 123 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|

مرتبسازی مبنایی یا پایه ای (یا به انگلیسی Radix sort) لیستی است با اندازه ثابت و اعضایی با طول k که در زمان $O(Kn)$ انجام می دهد. ابتدا از کم ارزشترین رقم در آرایه شروع کرده و الگوریتم را برای آن اجرا می کنیم، سپس به ارقامی با یک ارزش بیشتر می رویم و اینکار را ادامه می دهیم تا به با ارزش ترین رقم در میان اعداد آرایه برسیم. طریقه اجرای الگوریتم نیز در هر مرحله توضیح داده شده است. این الگوریتم از رده الگوریتم های خطی است که تابع رشد آن با حافظه و زمان رابطه خطی خود را حفظ می کند.

در این الگوریتم نیاز به دو آرایه با طول 10 که اعداد صفر تا نه را شامل می شود داریم.

برای پر کردن آرایه اول ابتدا به سراغ کم ارزشترین رقم در اعداد آرایه رفته و با توجه به آن رقم، اندیس آرایه شماره اول را یک واحد اضافه می کنیم، این کار را برای تمامی خانه های موجود در آرایه انجام می دهیم.

در مرحله بعدی برای پر کردن آرایه دوم اولین عدد را قرار داده و برای باقی اعداد، اعداد موجود در آرایه اول را از ابتدا به صورت جفتی جمع نموده و در اندیس آن قرار می دهیم، این کار را برای تمامی اعداد موجود در آرایه دوم نیز انجام می دهیم تا آرایه دوم نیز پر شود.

در مرحله سوم از آخر آرایه اصلی که اعداد را می خواهیم در آن مرتب کنیم شروع کرده و کم ارزشترین رقم آن که یکان است در ابتدا است را در نظر گرفته و اندیس آن عدد را در آرایه دوم می یابیم، از آن یک واحد کاسته و آن را در اندیس خانه آرایه جدیدی که قرار است جایگزین آرایه اصلی شود، می نویسیم، این کار را تا ابتدای آرایه برای تمامی عناصر انجام داده و مرحله اول مرتب سازی بر اساس پایه را انجام می دهیم.

حال با توجه به تعداد رقم های موجود در اعداد آرایه باید این مراحل را انجام دهیم .

در ادامه برای این مثال این عملیات بیان شده به ترتیب نشان داده شده است:

در این سوال تعداد ارقام بیشترین عدد موجود در جدول 3 عدد است پس نیاز داریم در 3 مرحله این عملیات را انجام دهیم:

مرحله اول :

| | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 324 | 239 | 879 | 358 | 225 | 137 | 317 | 416 | 123 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|

آرایه شماره اول را با استفاده از تعداد تکرار آن رقم در اعداد آرایه شکل می دهیم، برای مثال عدد یکان 317، عدد 7 است، پس به خانه اندیس شماره 7 یک واحد می افزاییم:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

در ماتریکس شماره دوم ابتدا رقم اول را در ابتدای ماتریکس قرار می دهیم و برای پر کردن بقیه ارقام آن عدد را رقم بعدی به صورت جفتی جمع کرده و در آن قرار می دهیم، عدد بعدی را با مجموع اعداد قبلی جمع کرده و در آن خانه قرار می دهیم، بدین صورت همه خانه های موجود در آرایه را پر می کنیم :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 2 | 3 | 4 | 6 | 7 | 9 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

در مرحله آخر نیز برای اعداد یکان آرایه تشکیل می دهیم، بدین صورت که از آخر آرایه اصلی شروع کرده و رقم یکان آن را جدا کرده و اندیس آن را در آرایه دوم یافته و آن را منفی یک کرده و خانه آرایه جدید قرار می دهیم :

| | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 324 | 225 | 416 | 137 | 317 | 358 | 239 | 879 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

اکنون این آرایه بر اساس یکان مرتب شده است، حال به ارزش رقم بعدی یعنی دهگان می روییم و آرایه را دوباره بر طبق همین منوال اما با آرایه مرتب شده ارزش یکان مرتب می کنیم :

آرایه شماره یک : تعداد تکرار رقم ها در ارزش دهگان، مقادیر عدد به ازای عدد 137 یک واحد به اندیس شماره 3 می افزاییم :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 3 | 2 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

آرایه شماره دو :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 5 | 7 | 7 | 8 | 8 | 9 | 9 | 9 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

و آرایه شماره سوم که مرتب شده آرایه ارزش یکان است:

| | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 416 | 317 | 123 | 324 | 225 | 137 | 239 | 358 | 879 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

در آخرین مرحله، همانند فرآیند قبلی، اینبار باید برای ارزش صدگان این عملیات را انجام دهیم، در این مرحله به دلیل 3 رقمی بودن همه اعداد موجود در آرایه، فرآیند مرتب سازی به اتمام می رسد و آرایه ای که به دست می آوریم، آرایه مرتب شده است :

آرایه شماره اول: این آرایه بر اساس تعداد تکرار ارقام در مرتبه ارزش دهگان تشکیل می شود :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 2 | 3 | 1 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

آرایه شماره دو: این آرایه نیز جفتی برای اعداد موجود در آرایه اول تشکیل می شود :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 4 | 7 | 8 | 8 | 8 | 8 | 9 | 9 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

و در آخرین مرحله، آرایه مرتب شده را می نویسیم :

| | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 123 | 137 | 225 | 239 | 314 | 324 | 358 | 416 | 879 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

بدین صورت این الگوریتم برای بقیه اعداد کار میکند، این الگوریتم از سری الگوریتم هایی که است که با الگوریتم های دیگر ترکیب شده و از لحاظ مرتبه زمانی و حافظه اجرایی بهینه تر شده و برای کاربرد های مختلف استفاده می شود.

2- در سناریوی زیر الگوریتم مرتب سازی ای که بهترین زمان اجرا را خواهد داشت انتخاب نموده و انتخاب خود را توضیح دهید .

فرض کنید صاحب یک سایت پر بیننده با تعداد زیادی کاربر هستید و برای تولید یک سری آمار، میخواهیم آن ها را بر اساس شناسه شان(id) که یک عدد صحیح 64 بیتی استف مرتب نمائید .

➤ مرتب سازی هرمی :

این مرتب سازی از خانواده مرتب سازی انتخابی است، بر اساس هرم دودویی کار میکند که به دوستانه کلی minheap و maxheap تقسیم می شود، درجا بوده و البته پایدار نیست، پیچیدگی زمانی آن از درجه $O(n \log n)$ است، در بدترین حالت این الگوریتم به پیچیدگی زمانی $O(n^2)$ می رسد که برای تعداد زیادی داده ناکارآمد محسوب می شود.

در اینجا نیز به دلیل وجود تعداد زیادی داده و اینکه این الگوریتم برای تعداد زیادی داده ضعف نشان می دهد، بهتر است که از این الگوریتم مرتب سازی استفاده نکنیم.

➤ مرتب سازی شمارشی :

این مرتب سازی از الگوریتم های خطی است که از نظر زمانی (و حافظه) از الگوریتم های دیگر بهینه تر هستند. در این مرتب سازی کمینه اعداد را می توانیم کمترین عدد موجود در آرایه در نظر بگیریم و بیشینه اعداد را بیشترین عدد موجود در آرایه و آرایه را به گونه ای تشکیل دهیم که به جای شروع از صفر تا بیشینه، از کمترین عدد موجود در آرایه شروع شود در اینجا به دلیل ذکر کردن 64 بیتی بودن داده ها، می توانیم این محدودیت را اعمال کنیم .

این الگوریتم پایدار با پیچیدگی زمانی $O(n+k)$ می باشد که در آن n طول آرایه ورودی و k طول آرایه شمارش است.

همچنین برای اینکه این الگوریتم کارآمد باشد لازم است k از مرتبه n باشد تا پیچیدگی نهایی آن از نوع $O(n)$ باشد.

➤ مرتب سازی درجی:

مرتب سازی با الگوریتم ساده ای است که در هر مرحله یک عنصر را در محل مناسبش در میان سایر داده ها که آن ها هم در مراحل مشابه مرتب شده اند، درج میکند.

این الگوریتم برای داده های مرتب بهترین عملکرد را از خود نشان می دهد، اما در بدترین حالت به پیچیدگی زمانی $O(n^2)$ می رسد که زمان زیادی محسوب می شود.

این الگوریتم معمولاً برای مرتب سازی آرایه هایی با اندازه کوچک استفاده می شود و در حجم زیاد آرایه از خود ضعف نشان می دهند.

➤ مرتب سازی مبنایی :

الگوریتمی از مرتبه الگوریتم های خطی است که لیستی از n داده با طول حداکثر k را در $O(nk)$ مرتب می کند. این الگوریتم بر اساس بارزشتترین یا کم ارزشترین رقم شروع به مرتب سازی میکند و برای بخشی از آن از الگوریتم مرتب سازی شمارشی استفاده می کند، روی کاغذ این الگوریتم به دلیل مطابقت با محدودیت های واقع شده در صورت سوال، بهترین نوع مرتب سازی برای این کاربرد است.

➤ مرتب سازی ادغامی :

این الگوریتم به روش تقسیم و حل لیست داده شده را مرتب می کنید و در هر مرحله آرایه را دو به دو به قسمت با اندازه تقریباً مساوی تقسیم کرده و هر زیرلیست را به صورت بازگشتی مرتب و سپس قسمت های مرتب شده را با یکدیگر ترکیب میکند.

پیچیدگی این الگوریتم از مرتبه $O(n \log n)$ می باشد، از نظرات مختلفی می توان این الگوریتم را به عنوان بهترین انتخاب برای مرتب سازی این لیست در نظر گرفت.

زیرا در این سوال هیچ دو شناسه ای با هم برابر نیست و می توان آن را در پردازنده های چند هسته ای به صورت موازی پیاده سازی کرد.

○ بنده تمامی الگوریتم ها را در زبان سی پلاس پلاس پیاده سازی کردم و لیست های مختلفی را به عنوان نمونه به آن دادم و زمان آن ها را اندازه گیری کردم تا از نتیجه بدست آمده اطمینان حاصل کنم، به دلیل زیاد بودن تعداد ارقام، مرتب سازی مبنایی ضعف زیادی از خود نشان داد، مرتب سازی ادغامی تا تعدادی از داده ها از مرتب سازی شمارشی بهتر عمل میکند ولی در حجم بسیار زیادی داده با تعداد ارقام زیاد (که در اینجا تقریباً $16 = 64/4$ رقم با در نظر گرفتن 4 بیتی بودن هر رقم) مرتب سازی شمارشی از بقیه مرتب سازی ها پیشی گرفته و این اختلاف بسته به تعداد داده ها بیشتر و بیشتر می شود و با فاصله نسبتاً زیادی از بقیه الگوریتم ها در این مورد بهتر کار میکند.