به نام خدا

ساختمان داده

آرش شفیعی



مرتبسازي

مرتبسازي

ساختمان داده

در این بخش در مورد چند الگوریتم مرتبسازی صحبت خواهیم کرد و داده ساختارهای مورد نیاز برای این الگوریتمها را معرفی خواهیم کرد.

- $^{-}$ مسئله مرتبسازی 1 به شرح زیر است :
- $\langle a_1, a_2, \cdots, a_n \rangle$ ورودی : دنبالهای از n عدد به صورت –
- خروجی : ترتیبی (یا جایگشتی) از ورودیها به صورت $\langle a_1',a_2',\cdots,a_n' \rangle$ به طوری که $a_1' \leqslant a_2' \leqslant \cdots \leqslant a_n'$
- دنباله ورودی معمولا به صورت یک آرایه n عنصری نشان داده می شود، گرچه می تواند با استفاده از داده ساختارهای دیگری مانند لیست پیوندی نیز مدلسازی شود.

¹ sorting problem

- معمولا در عمل عناصر یک آرایه مقادیر مجزا نیستند، بلکه هرکدام نماینده یا مجموعه ای از داده ها هستند که آن را رکورد 1 مینامیم.
- یک رکورد مجموعهای از دادهها با نوعهای دادهای متفاوت است که با یک کلید مشخص می شود. برای مثال اطلاعات یک دانشجو یک رکورد را تشکیل می دهد و شماره دانشجویی کلید رکورد است.
 - در عمل وقتی کلیدهای تعدادی رکورد مرتب میشوند، دادههای رکوردهای آنها هم به همراه کلیدها جابجا میشوند. اگر اطلاعات یک رکورد زیاد باشند میتوان در کنار کلید اشاره گری به اطلاعات رکورد مربوط به کلید ذخیره کرد و اشاره گر را به همراه کلید در رکورد جابجا نمود.
- در این قسمت تنها در مورد الگوریتمهای مرتبسازی صحبت می کنیم صرف نظر از اینکه چه نوع اطلاعاتی را مرتب می کنیم. البته برای توصیف الگوریتمها فرض می کنیم ورودی از اعداد تشکیل شده است.

¹ record

- بسیاری از دانشمندان علوم کامپیوتر بر این باورند که مرتبسازی یکی از پایه ای ترین مسائل در مطالعه الگوریتم هاست. دلایل زیادی موجب این اهمیت است از جمله این که:
- بسیاری از برنامهها نیاز به مرتبسازی اطلاعات دارند. برای مثال، در یک سامانه دانشگاهی نیاز داریم دانشجویان را بر اساس یک شاخص مثلاً شماره دانشجویی و یا نامخانوادگی مرتب کنیم.
- بسیاری از الگوریتمهای کامپیوتری نیاز به مرتبسازی دارند. برای مثال، یک برنامه کامپیوتری که اشیاء گرافیکی را پردازش می کند برای به تصویر کشیدن اشیاء نیاز دارد آنها را با اساس فاصلهٔ آنها از بیننده تصویر م تب کند.

- میتوانیم روشها و رویکردهای مختلف الگوریتمی را توسط الگوریتمهای مرتبسازی مورد بررسی قرار دهیم.
 در طی سالیان، الگوریتمهای مرتبسازی مختلفی با رویکردها و روشهای گوناگون به وجود آمدهاند که
 مطالعه هرکدام به فهم بهتر مبحث الگوریتمها و دادهساختارها کمک میکند.
- میتوانیم روشهای مختلف مرتبسازی را براساس عملکردشان با یکدیگر مقایسه کنیم. با توجه به پراکندگی و ساختار ورودی مورد نظر برای جستجو، الگوریتم مناسب را انتخاب کنیم. برای هریک از روشهای مرتبسازی میتوانیم حد آستانه زمان اجرا را محاسبه و اثبات و با الگوریتمهای دیگر مقایسه کنیم.

در مقدمه تحلیل الگوریتمها مرتبسازی درجی 1 را معرفی کردیم که در بدترین حالت یک آرایه n عنصری را در زمان $\Theta(n^2)$ مرتب می کند. این الگوریتم به فضای اضافی برای مرتبسازی نیاز ندارد و آرایه را درجا 2 مرتب می کند که می تواند برای آرایههای بسیار بزرگ یک مزیت محسوب شود.

¹ insertion sort

² in place

در این بخش درمورد یک الگوریتم مرتبسازی دیگر صحبت خواهیم کرد.

- مرتب میکند. $O(n \lg n)$ مرتب میکند. $O(n \lg n)$ مرتب میکند.
- مرتبسازی هرمی از یک داده ساختار مهم به نام هرم یا هیپ 2 استفاده میکند که به معرفی این داده ساختار و موارد استفاده از آن خواهیم پرداخت.

ساختمان داده مرتبسازی ۷ / ۷

¹ heap sort

² heap

مرتبسازی هرمی و مرتبسازی درجی در دسته ای از الگوریتمهای مرتبسازی به نام مرتبسازیهای مقایسه ای مقایسه ای قرار می گیرند. دو الگوریتم مرتبسازی مقایسه ای مهم دیگر که در مبحث طراحی الگوریتمها به آن پرداخته می شود مرتبسازی ادغامی 1 و مرتبسازی سریع 2 هستند. می توان اثبات کرد که هر نوع مرتبسازی مقایسه ای نمی تواند در زمان کمتر از $n \lg n$ انجام شود بنابراین کران پایین زمان اجرای مرتبسازی های مقایسه ای $\Omega(n \lg n)$ است.

V° / A

¹ merge sort

² quick sort

- اگر مجموعه ای که می خواهیم مرتب کنیم شامل اعداد حسابی $\{0,1,\cdots,k\}$ باشد با استفاده از یک حافظه اضافی می توانیم زمان اجرا را به $\{0,k+n\}$ کاهش دهیم. این الگوریتم مرتبسازی را الگوریتم مرتبسازی شمارشی $\{0,k+n\}$ می نامیم. اگر $\{0,k+n\}$ باشد، مرتبسازی شمارشی اعداد را در زمان خطی نسبت به تعداد اعداد آرایه مرتب می کند.

الگوریتم مرتبسازی سطلی 2 الگوریتم دیگری است که در دسته الگوریتمهای مرتبسازی خطی قرار می گیرد. اگر اطلاعاتی از توزیع احتمالی اعداد در آرایه ورودی داشته باشیم و بدانیم $^{\rm n}$ عدد ورودی دارای توزیع احتمالی یکنواخت $^{\rm 8}$ هستند می توانیم از مرتبسازی سطلی استفاده کنیم که در حالت میانگین اعداد را در زمان ${\rm O}({\rm n})$ مرتب می کند.

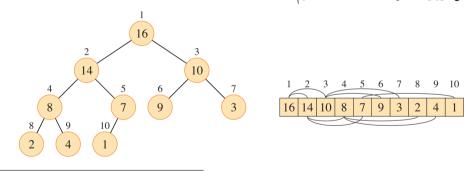
¹ counting sort

² bucket sort

³ uniform distribution

- داده ساختار هرم 1 یا هرم دودویی 2 آرایه ای از اشیاء است که میتوانیم آن را به صورت یک درخت دودویی نشان دهیم.

- در شکل زیر یک نمونه از داده ساختار هرم نشان داده شده است.



¹ heap

Vo / \o

² binary heap

- هر رأس در درخت متناظر با یکی از عناصر آرایه است. درخت در همهٔ سطوح به غیر از سطح آخر یا سطح برگها کامل است و در سطح آخر از چپ به راست پر شده است.
- آرایه A.heap-size که یک هرم را نمایش می دهد یک شیء با ویژگی A.heap-size است که تعداد عناصر هرم را که در آرایه A.in خخیره شده است نمایش می دهد. بنابراین اگرچه A[1:n] می تواند از اعدادی تشکیل شده باشد، تنها عناصر A.heap-size عناصر معتبر باشد، تنها عناصر A.heap-size عناصر معتبر هرم هستند. اگر A.heap-size باشد، هرم خالی است.
- ریشه درخت A[1] است و به ازای اندیس i از یک رأس در درخت، روش سادهای برای محاسبه اندیسهای پدر، فرزند چپ و فرزند راست آن رأس وجود دارد.

الگوریتمهای مرتبسازی - در زیر روش محاسبه اندیسهای پدر، فرزند چپ و فرزند راست نشان داده شده است.

Algorithm Parent

function PARENT(i)

1: return | i/2 |

Algorithm Left

function LEFT(i)

1: return 2i

Algorithm Right

function RIGHT(i)

1: return 2i + 1

- در بسیاری از کامپیوترها مقدار 2i میتواند با یک دستور ماشین توسط انتقال عدد دودویی متناظر با i یک واحد به سمت چپ محاسبه شود. همچنین مقدار i/2 توسط یک انتقال به راست محاسبه می شود.

- $^{-}$ دو نوع هرم دودویی وجود دارد : هرم بیشینه 1 و هرم کمینه 2 .
 - هر دو نوع هرم دارای ویژگی هرم 3 هستند.
- در هرم بیشینه ویژگی هرم بیشینه 4 بدین صورت است که به ازای هر رأس i به غیر از ریشه داریم :

 $A[Parent(i)] \geqslant A[i]$

¹ max heap

² min heap

³ heap property

⁴ max heap property

- این ویژگی بدین معناست که مقدار یک رأس در درخت حداکثر برابر با مقدار رأس پدر آن رأس است . بنابراین بزرگترین عنصر در یک هرم بیشینه در ریشه ذخیره می شود و یک زیردرخت دارای یک ریشه مقداری بزرگتر از ریشه را شامل نمی شود.

- یک هرم کمینه به روش عکس هرم بیشینه ذخیرهسازی را انجام میدهد. ویژگی هرم کمینه 1 بدین صورت است که به ازای هر رأس i به غیر از ریشه، داریم:

 $A[Parent(i)] \leq A[i]$

- بنابراین دریک هرم کمینه کوچکترین عنصر در ریشه ذخیره میشود.

ساختمان داده مرتبسازی ۲۰ / ۲۷

¹ min heap property

- الگوريتم مرتبسازي هرمي از يک هرم بيشينه استفاده مي کند که بعدها به آن خواهيم پرداخت.

- هرم کمینه برای پیادهسازی صف اولویت 1 استفاده می شود که درمورد آن نیز بعدها صحبت خواهیم کرد.

¹ priority queue

- اگر یک هرم را به صورت یک درخت نمایش دهیم، میتوانیم ارتفاع 1 یک رأس از درخت را تعریف کنیم.

- ارتفاع یک رأس از درخت، تعداد یالهایی است که برروی بلندترین مسیر ساده از یکی از برگهای درخت به آن رأس قرار دارد. ارتفاع یک هرم ارتفاع ریشه هرم است.

است. ارتفاع آن $\Theta(\lg n)$ است، ارتفاع آن $\Theta(\lg n)$ است.

- عملیات اصلی که بر روی هرم اجرا میشوند حداکثر در زمانی متناسب با ارتفاع هرم اجرا میشوند و در نتیجه حداکثر به زمان $O(\lg n)$ نیاز دارند.

ساختمان داده مرتبسازی ۷۰/۱۶

¹ height

² complete binary tree

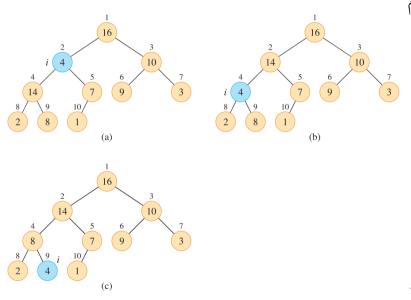
تابع Max-Heapify تابعی است که ویژگی هرم بیشینه را حفظ میکند. ورودی آن یک آرایه A با اندازه heap-size و یک اندیس i در آرایه است. وقتی این تابع فراخوانی می شود، تابع فرض را بر این می گذارد که در درخت دودویی با ریشههای i Left(i) و i Right(i) هرم بیشینه هستند، اما i ممکن است کوچکتر از فرزندانش باشد. تابع i Max-Heapify مقدار i را در هرم بیشینه پایین می برد تا جایی که ویژگی هرم سشینه به دست آید.

Algorithm Max Heapify

```
function Max-Heapify(A.i)
1: 1 = I.eft(i)
2: r = Right(i)
3: if 1 \leq A.heap-size and A[1] > A[i] then
4:
      largest = 1
5: else largest = i
6: if r \leq A.heap-size and A[r] > A[largest] then
7:
      largest = r
8: if largest \neq i then
      exchange A[i] with A[largest]
9:
      Max-Heapify (A, largest)
10:
```

- شکل زیر عملیات Max-Heapify را نشان میدهد. هرگام بیشترین مقدار عناصر A[i] و A[i] و A[i] A[i] و A[i] A[i] A[i] را تعیین میکند و اندیس بزرگترین عنصر را در متغیر largest ذخیره میکند. اگر A[i] بزرگترین عنصر است، زیر درخت با ریشه i یک هرم بیشینه است و نیاز به انجام هیچ عملیاتی نیست. در غیر اینصورت، یکی از دو فرزند دارای بیشترین مقدار است و مکان i و targest تعویض میشود که باعث میشود رأس i و فرزندانش ویژگی هرم بیشینه را حفظ کنند. اما اکنون زیر درخت با رأس largest ممکن است ویژگی هرم بیشینه را نقض کند، پس دوباره تابع Max-Heapify به صورت بازگشتی باید فراخوانی شده.

حفظ ویژگی هرم



٧٠ / ٢٠

مرتبسازي

ساختمان داده

- برای تحلیل تابع Max-Heapify ، فرض کنید T(n) زمان اجرا در بدترین حالت باشد که تابع برای یک زیر درخت با اندازه n صرف می کند.
- برابی با A[Right(i)] و A[Left(i)] ، A[i] برابر با A[i] برابر با A[Right(i)] و A[Ieft(i)] برابر با O(1) است. تابع به صورت بازگشتی برای یکی از زیردرختها فراخوانی می شود. هرکدام از زیردرختها حداکثر 2n/3 رأس دارند و بنابراین می توانیم زمان اجرای تابع Max—Heapify را با رابطه بازگشتی زیر توصیف کنیم.

$$T(n) \leq T(2n/3) + \Theta(1)$$

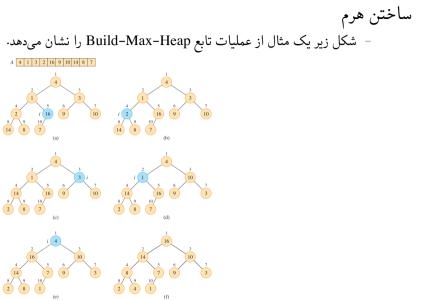
- $\mathsf{T}(n) = O(\lg n)$ با حل کردن این رابطه بازگشتی توسط قضیه اصلی به دست می آوریم
- همچنین میتوانیم بگوییم زمان اجرای تابع Max-Heapify بر روی رأسی به ارتفاع h برابر با (O(h است.

- تابع Build-Max-Heap آرایه [1: n] را به یک هرم بیشینه تبدیل می کند.
 - عناصر زیر آرایه A[|n/2|+1:n] همه برگهای درخت هستند.
- تابع Build-Max-Heapify به ازای همه عناصر غیر برگ تابع Max-Heapify را فراخوانی می کند.

Algorithm Build Max Heap

function BUILD-MAX-HEAP(A,n)

- 1: A.heap-size = n
- 2: for $i = \lfloor n/2 \rfloor$ downto 1 do
- 3: Max-Heapify(A,i)



- برای اینکه نشان دهیم چرا تابع Build-Max-Heap درست کار میکند، باید اثبات کنیم در ابتدای حلقه $i+1,i+2,\cdots,n$ هر یک از رأسهای for
- باید نشان دهیم این گزاره در ابتدای اولین تکرار حلقه درست است و در هر تکرار حلقه اگر گزاره در ابتدای یک تکرار درست باشد، در ابتدای تکرار بعدی نیز درست است. بدین ترتیب به روش استقرایی اثبات میکنیم الگوریتم درست عمل میکند.

- پایهٔ استقرا : قبل از اولین تکرار حلقه، $\lfloor n/2 \rfloor = i$. هریک از رئوس بایهٔ است و بنابراین به طور بدیهی ریشه یک هرم بیشینه است. $\lfloor n/2 \rfloor + 1, \lfloor n/2 \rfloor + 2, \cdots, n$
- گام استقرا : اندیس فرزندان رأس i بزرگتر از رأس i است. طبق فرض استقرا هردوی فرزندان ریشه هرم بیشینه هستند. این دقیقا شرایطی است که برای فراخوانی تابع Build-Max-Heap(A،i) نیاز داریم تا رأس i را در ریشه هرم بیشینه قرار دهیم. همچنین فراخوانی تابع Max-Heapify ویژگی هرم را حفظ می کند به طوری که رئوس $i+1,i+2,\cdots,n$ همه ریشه هرم بیشینه خواهند بود. بنابراین در تکرار بعدی حلقه این رئوس هریک ریشه یک هرم بیشینه خواهند بود.
- خاتمهپذیری: حلقه دقیقا $\lfloor n/2 \rfloor = i$ بار تکرار می شود بنابراین قطعا الگوریتم خاتمه می یابد. در پایان i = 0 است بنابراین طبق اثبات استقرایی هریک از رئوس i = 0 ریشه یک هرم بیشینه هستند و رأس 1 ریشه هرم بیشینه به دست آمده است.

- میتوانیم یک کران بالای ساده بر روی زمان Build-Max-Heap به صورت زیر محاسبه کنیم. زمان اجرای هر فراخوانی O(n) Build-Max-Heap تعداد $O(\lg n)$ Max-Heapify هر فراخوانی انجام میدهد. بنابراین زمان اجرا $O(n \lg n)$ است. این کران بالا اگرچه درست است، اما به میزان کافی دقیق نیست.
 - میتوانیم یک کران مجانبی دقیق تر برای این تابع محاسبه کنیم. مشاهده می کنیم که زمان اجرای Max-Heapify بستگی به ارتفاع درختی دارد که برروی آن اجرا می شود. در یک هرم شامل n عنصر ارتفاع $\lfloor \lg n \rfloor$ است و حداکثر تعداد $\lfloor n/2^{h+1} \rfloor$ رأس با ارتفاع n وجود دارد.
 - زمان مورد نیاز برای تابع Max-Heapify وقتی بر روی رأسی با ارتفاع h فراخوانی میشود برابراست با .O(h)

وا به Build-Max-Heap را به انگاه زمان اجرای طریب ثابت در تحلیل مجانبی باشد، آنگاه زمان اجرای صورت زیر می توانیم توصیف کنیم.

$$\sum_{h=0}^{\lfloor \lg n \rfloor} \lceil \frac{n}{2^{h+1}} \rceil ch$$

 $\lfloor n/2^{h+1} \rfloor \geqslant 1/2$ به ازای $0 \leqslant h \leqslant \lfloor \lg n \rfloor$ داریم –

$$\lceil n/2^{h+1} \rceil \leqslant n/2^h$$
، بنابراین $\lceil c \rceil \leqslant 2x$ داریم $x \geqslant 1/2$ داریم که به ازای که به ازای حمینین از آنجایی که به ازای حمینین از آنجایی که به ازای $x \geqslant 1/2$

- بنابراین میتوانیم رابطه زیر را به دست آوریم.

$$\sum_{k=0}^{\lfloor \lg n \rfloor} \lceil \frac{n}{2^{h+1}} \rceil ch \leqslant cn \frac{1/2}{(1-1/2)^2} = O(n)$$

- بنابراین میتوانیم یک هرم بیشینه از یک آرایه غیر مرتب در زمان خطی بسازیم.
- برای ساختن هرم کمینه از تابع Build-Min-Heap استفاده میکنیم که شبیه Build-Max-Heap است با این تفاوت که تابع Min-Heapity فراخوانی می شود.

الگوریتم مرتبسازی هرمی 1 ، که با تابع Heapsort در زیر نشان داده شده است با فراخوانی تابع Build-Max-Heap یک هرم بیشینه بر روی آرایه ورودی A[1:n] میسازد. از آنجایی که مقدار بیشینه عناصر در ریشه A[1] ذخیره شده است، الگوریتم مرتبسازی هرمی میتواند این مقدار بیشینه را در مکان درست آن یعنی A[n] قرار دهد.

حال اگر تابع رأس n را از هرم بردارید یا به عبارت دیگر A.heap-size را یک واحد کاهش دهد، فرزندان ریشه همچنان هرم بیشینه باقی می مانند. برای حفظ ویژگی هرم تابع Max-Heapify(A، ۱) فراخوانی می شود و در نتیجه A[1:n-1] یک هرم بیشینه می شود. این عملیات ادامه می یابد تا جایی که اندازه هرم برابر با ۲ شود.

ساختمان داده مرتبسازی ۲۹ / ۷۰

¹ heapsort algorithm

Algorithm Heapsort

function HEAPSORT(A,n)

1: Build-Max-Heap (A,n)

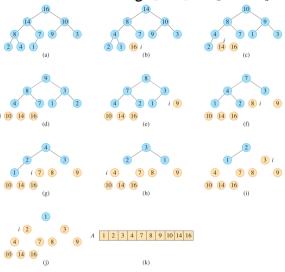
2: for i = n downto 2 do

3: exchange A[1] with A[i]

4: A.heap-size = A.heap-size-1

5: Max-Heapify (A,1)

- شکل زیر یک مثال از عملیات Heapsort را نشان میدهد.



V0 / W1

مرتبسازي

ساختمان داده

به زمان Build-Max-Heap اجرا می شود، زیرا فراخوانی تابع $O(n \lg n)$ به زمان $O(\lg n)$ به زمان $O(\lg n)$ نیاز دارد و هریک از n-1 فراخوانی تابع O(n)

صف اولویت

- داده ساختار هرم کاربردهای زیادی در الگوریتمهای مختلف دارد. یکی از کاربردهای آن صف اولویت 1 است.

- در اینجا درمورد یکی از انواع صف اولویت که صف اولویت بیشینه 2 است صحبت میکنیم که براساس هرم بیشینه است. صف اولویت کمینه 3 شبیه صف اولویت بیشینه است که از هرم کمینه استفاده میکند.

 4 یک صف اولویت داده ساختاری است برای نگهداری مجموعه 8 از عناصری که هرکدام با یک کلید 4 مشخص شدهاند. صف اولویت بیشینه عملیات زیر را پشتیبانی میکند.

ساختمان داده مرتبسازی ۷۰ / ۳۳

¹ priority queue

² max priority queue

³ min priority queue

⁴ kev

- عملیات درج (S،x،k) عنصر x با کلید : Insert(S،x،k) عملیات درج عملیات درج $S = S \cup \{x\}$
- عملیات بیشینه (Maximum(S : عنصری از S که بزرگترین کلید را دارد را بر میگرداند.
- عملیات استخراج بیشینه (Extract-Max(S : عنصری از S که بزرگترین کلید را دارد را از مجموعه حذف کرد، و باز می گرداند.
 - میدهد Increase–key(S(x) افزایش میدهد نصر x را با مقدار جدید x افزایش میدهد فرض بر این است که مقدار x بزرگتر یا مساوی مقدار فعلی کلید x است.

صف اولويت

- صف اولویت می تواند کاربردهای زیادی داشته باشد. برای مثال واحدهای کاری 1 در یک سیستم عامل به هنگام اجرا وارد یک صف می شوند، اما یک زمانبندی ممکن است بخواهد آنها را براساس اولویت شان از صف خارج کند که در اینجا می توانیم از یک صف اولویت بیشینه استفاده کنیم. هرگاه یک واحد کاری به اتمام رسید، زمانبند 2 واحد کاری با بیشترین اولویت را توسط Extract-Max از صف اولویت خارج می کند. هرگاه یک واحد کاری جدید توسط کاربر تعریف شد، توسط Insert وارد صف اولویت می شود.
 - یک صف اولویت کمینه نیز عملیات Extract-Min ، Minimum ، Insert و Decrease-Key را یشتبانی می کند.
 - ممکن است یک زمانبند بخواهد واحدهای کاری که کمترین زمان اجرا دارند زودتر اجرا کند. در این صورت کلید عناصر زمان اجرای مورد نیاز آنهاست و از یک صف اولویت کمینه استفاده می شود.

ساختمان داده مرتبسازی ۷۵ / ۷۵

¹ jobs or tasks

² Scheduler

صف اولويت

و تابع Max-Heap-Maximum عملیات Max-Heap-Maximum و تابع $\Theta(1)$ و تابع Max-Heap-Extract-Max عملیات M عملیات M

Algorithm Max Heap Maximum

function Max-Heap-Maximum(A)

- 1: if A.heap-size < 1 then
- 2: error "heap underflow"
- 3: return A[1]

Algorithm Max Heap Extract Max

function MAX-HEAP-EXTRACT-MAX(A)

- 1: max = Max-Heap-Extract-Max (A)
- 2: A[1] = A[A.heap-size]
- 3: A.heap-size = A.heap-size 1
- 4: Max-Heapify(A,1)
- 5. return max

- در این الگوریتمها فرض میکنیم Max-Heapify عناصر صف اولویت را بر اساس ویژگی کلید آنها مقایسه میکند و همچنین وقتی دو عنصر در آرایه جابجا میشوند، درواقع اشاره گری به عناصر آنها جابجا میشود.

رمان اجرای Max-Heap-Extract-Max برابر با $O(\lg n)$ است، زیرا تنها تعدادی عملیات در زمان Max-Heapify انجام میدهد.

- تابع Max-Heap-Increase-Key عمليات Max-Heap-Increase-Key را انجام مي دهد.

Algorithm Max Heap Increase Key

function Max-Heap-Increase-Key(A,x,k)

1: if k < x.key then

2: error "new key is smaller than current key"

3: x.key = k

4: find the index i in array A where object x occurs

5: while i > 1 and A[Parent(i)].key < A[i].key do

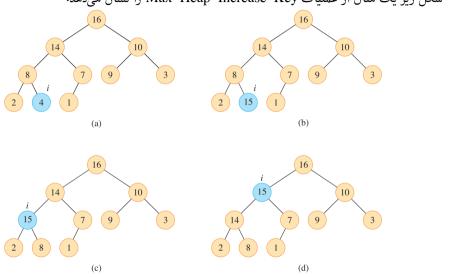
6: exchange A[i] with A[Parent(i)], updating the information that priority queue objects to array indices

7: i = Parent(i)

- این تابع ابتدا بررسی میکند که کلید جدید K کلید فعلی x را کاهش نمیدهد و در صورتی که مشکل وجود نداشت، مقدار کلید x را تغییر میدهد. سپس این تابع اندیس i از x را در آرایه پیدا میکند به طوری که A[i] داد با x باشد.
- از آنجایی که افزایش کلید A[i] ممکن است ویژگی هرم بیشینه را نقض کند، تابع یک مسیر ساده از رأس به سمت ریشه طی می کند تا مکان مناسب رأس را با کلید جدید پیدا کند.
- وقتی تابع Max-Heap-Increase-Key این مسیر را طی میکند، به طور مکرر کلید یک عنصر را با کلید پدر آن مقایسه میکند و در صورتی که کلید عنصر بزرگتر بود، اشاره گرهای آنها با جابجا میکند و در صورتی که کلید عنصر از کلید پدر کوچکتر بود، الگوریتم به پایان میرسد، زیرا ویژگی هرم بیشینه برقرار خواهد بود.

صف اولويت

- شکل زیر یک مثال از عملیات Max-Heap-Increase-Key را نشان میدهد.



Yo / 40

مرتبسازي

ساختمان داده

- زمان اجرای عملیات افزایش کلید Max-Heap-Increase-Key برای هر م با n عنصر برابر با $O(\lg n)$ است زیرا مسیری که رأس تا ریشه طی می کند $O(\lg n)$ است.

- تابع Max-Heap-Insert عملیات درج Insert را پیادهسازی میکند.

Algorithm Max Heap Insert

function Max-Heap-Insert(A,x,n)

1: if A.heap-size == n then

2: error "heap overflow"

3: A.heap-size = A.heap-size + 1

4: k = x.key

5: $x.key = -\infty$

6: A[A.heap-size] = x

7: map x to index heap-size in the array

8: Max-Heap-Increase-Key (A,x,k)

صف اولویت

x این تابع به عنوان ورودی آرایه x را که یک هرم بیشینه را پیادهسازی کرده است، به همراه عنصر جدید x دریافت میکند و آن را در آرایه x با اندازه x درج میکند. این تابع بررسی میکند مکان کافی در آرایه برای درج عنصر جدید وجود دارد. سپس یک برگ با کلید x به هرم اضافه میکند. سپس تابع درج عنصر جدید وجود دارد. سپس یک برگ با کلید تا کلید عنصر اخیرا افزوده شده را تنظیم و آن را در مکان مناسب خود قرار دهد تا ویژگی هرم برقرار شود.

است. $O(\lg n)$ بر روی یک هرم با n عنصر برابر با $O(\lg n)$ است.

مرتبسازی در زمان خطی

- الگوریتمهای مرتبسازی بهینه یک آرایه n تایی را در زمان $\mathrm{O}(n \lg n)$ مرتب میکنند.
- برای مثال الگوریتمهای مرتبسازی هرمی و ادغامی در بدترین حالت و مرتبسازی سریع در زمان متوسط مرتبسازی n عنصر را در زمان $O(n \lg n)$ انجام میدهند.
 - این الگوریتمها که با مقایسه عناصر آرایه مرتبسازی را انجام میدهند در دسته الگوریتمهای مرتبسازی مقایسهای ¹ قرار می گیرند.
- در این قسمت در مورد چند الگوریتم مرتبسازی صحبت میکنیم که آرایه را در زمان خطی نسبت به تعداد عناصر آرایه مرتب میکنند.

weight الله مرتبسازي مرتبسازي ۷۰/۴۴

¹ comparison sort

- ابتدا اثبات می کنیم که هر الگوریتم مرتبسازی مقایسه ای کران پایین زمان اجرای $\Omega(n \lg n)$ در بدترین حالت دارد. به عبارت دیگر هیچ الگوریتم مقایسه ای وجود ندارد که آرایه n عنصری را در زمانی کمتر از $n \lg n$ بتواند مرتب کند.
- یک مرتبسازی مقایسه ای تنها با مقایسه عناصر دنباله $\langle a_1, a_2, \cdots, a_n \rangle$ در مورد ترتیب عناصر آن اطلاعات کسب می کند. بدین معنی که به ازای دو عنصر a_i و a_i این الگوریتمها یکی از مقایسههای اطلاعات کسب می کند. $a_i \geqslant a_j$ ، $a_i \geqslant a_j$ ، $a_i \geqslant a_j$ ، $a_i \leqslant a_j$ ، $a_i \leqslant a_j$ ، $a_i \leqslant a_j$ تعین کنند.
- برای اثبات کران پایین 1 مرتبسازیهای مقایسهای بدون از دست دادن عمومیت اثبات 2 ، فرض میکنیم عناصر ورودی متمایز هستند. بنابراین فرض میکنیم همهٔ مقایسهها به صورت $lpha_{
 m i} \leqslant lpha_{
 m i}$ هستند.

¹ lower bound

² without loss of generality

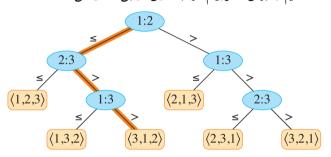
میتوانیم مرتبسازی مقایسهای را به صورت درختهای تصمیم 1 مدلسازی کنیم.

- یک درخت تصمیم یک درخت دودویی پر 2 است (بدین معنی که هر رأس یا یک برگ است و یا دو فرزند دارد) که مقایسههای بین عناصر آرایه را نمایش می دهد که توسط یکی از انواع مرتبسازی های مقایسه ای بر روی عناصر ورودی انجام شده است.

¹ decision trees

² full binary tree

- شکل زیر یک درخت تصمیم را برای الگوریتم مرتبسازی درجی نشان میدهد.



- n هر یک از رئوس میانی درخت تصمیم با i:j به ازای $i\geqslant 1$ و $i\geqslant i$ نمایش داده شدهاند به طوری که i تعداد عناصر دنباله ورودی است.
 - هر یک از برگها با یک جایگشت $\langle \Pi(1),\Pi(2),\cdots,\Pi(n) \rangle$ نمایش داده شدهاند.
 - اندیسهایی که در رئوس درخت تصمیم نوشته شده اند مکان اصلی عناصر را نشان میدهند اجرای درخت تصمیم بر روی یک ورودی متناظر با یک مسیر ساده از ریشه به یکی از برگهاست.
- $a_i \leqslant a_j$ هر رأس میانی یک مقایسه $a_i \leqslant a_j$ انجام میدهد. زیر درخت سمت چپ شرایطی را بررسی میکند که در آن $a_i \leqslant a_j$ است. با رسیدن به یک $a_i \leqslant a_j$ است. با رسیدن به یک $a_i \leqslant a_j$ تعیین شده است. ریشه ترتیب $a_{\Pi(n)} \leqslant \alpha_{\Pi(2)} \leqslant \cdots \leqslant \alpha_{\Pi(n)}$ تعیین شده است.
 - هر یک از n! جایگشت n عنصر باید در یکی از برگهای درخت تصمیم ظاهر شود.

كران پايين مرتبسازي مقايسهاي

- طول بلندترین مسیر بین ریشه و برگها زمان اجرای الگوریتم مرتبسازی مقایسهای در بدترین حالت را تعیین میکند. بنابراین بیشترین تعداد مقایسهها در یک مرتبسازی مقایسهای برابر با ارتفاع درخت تصمیم
 - کران پایین برروی ارتفاع همه درختهای تصمیم که در آنها یک ترتیب در برگهای ظاهر میشود، درواقع کران پایین برروی زمان اجرای هرنوع مرتبسازی مقایسه ای است.

- قضیه : هر الگوریتم مرتبسازی مقایسه ای نیاز به حداقل $\Omega(n \lg n)$ مقایسه در بدترین حالت برای مرتبسازی n عنصر دارد.
- اثبات: مرتبسازی n عنصر در یک الگوریتم مقایسه ای متناظر با یک درخت تصمیم با ارتفاع n و l برگ است. از آنجایی که ای یا چند برگ ظاهر می شوند $n! \leqslant l$ است. از آنجایی که یک درخت دودویی با طول n بیشتر از n! برگ ندارد، داریم:

$$n! \le l \le 2^h$$

- $h\geqslant\lg(n!)=\Omega(n\lg n)$ با گرفتن لگاریتم از دو طرف داریم
 - بنابراین مرتبسازی هرمی یک الگوریتم مرتبسازی بهینه است.

- مرتبسازی شمارشی 1 فرض میکند که هریک از n عنصر ورودی یک عدد صحیح در بازه O تا K است. این الگوریتم در زمان $\Theta(n+K)$ اجرا میشود، بنابراین وقتی K=O(n) باشد، مرتبسازی شمارشی در زمان $\Theta(n)$ اجرا میشود.
- مرتبسازی شمارشی ابتدا به ازای هر عنصر x تعداد عناصر کوچکتر یا مساوی x را تعیین میکند. سپس از این اطلاعات استفاده میکند تا مکان عنصر x را در آرایه خروجی تعیین کند.
- برای مثال اگر تعداد ۱۷ عنصر کوچکتر یا مساوی x هستند، مکان x در آرایه خروجی مکان ۱۷ است. همچنین باید شرایطی که چند عنصر میتوانند مساوی باشند را در نظر بگیریم زیرا نمیخواهیم همه عناصری که مقدارشان مساوی یکدیگر است را در یک مکان حافظه قرار دهیم.

ساختمان داده مرتبسازی ۷۰/۵۱

¹ countiong sort

مرتبسازی شمارشی

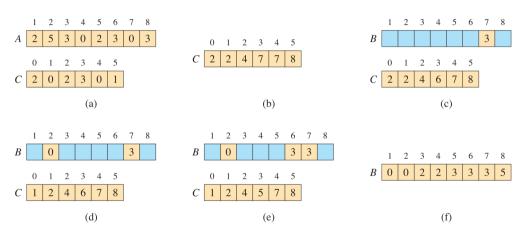
تابع Counting–Sort آرایه A[1:n] را که اندازه آن n است به همراه مقدار حدی k که یک عدد صحیح غیر منفی است را دریافت می کند و آرایه B[1:n] شامل عناصر مرتبشده را باز می گرداند. در این مرتبسازی از آرایه C[0:k] به عنوان فضای کمکی استفاده می کند.

Algorithm Counting Sort

```
function COUNTING-SORT(A, n, k)
1: let B[1:n] and C[0:k] be new arrays
2: for i = 0 to k do
3: C[i] = 0
4: for j = 1 to n do
5: C[A[i]] = C[A[i]] + 1
6: \triangleright C[i] now contains the number of elements equal to i.
7. for i = 1 to k do
8: C[i] = C[i] + C[i-1]
9: \triangleright C[i] now contains the number of elements less than or equal to i.
10: ▷ Copy A to B, starting from the end of A.
11: for j = n downto 1 do
12: B[C[A[j]]] = A[j]
13: C[A[j]] = C[A[j]] - 1 \triangleright to handle duplicate values
14: return B
```

م تبسازی

- شكل زير نحوه اجراى الگوريتم مرتبسازي شمارشي را نشان ميدهد.



Vo / D4

مرتبسا

ساختمان داده

- A بعد از حلقه خطوط ۲ و ۳ که آرایه C را با صفر مقداردهی اولیه می کند، حلقه خطوط ۴ و ۵ یک بار آرایه C را بررسی می کند تا اطلاعاتی درمورد آرایه به دست آورد. هربار یک عنصر ورودی که مقدار C[i] است پیدا می شود مقدار C[i] یک واحد افزایش پیدا می کند. بنابراین بعد از خط ۵، مقدار C[i] تعداد عناصر ورودی است که مقدارشان برابر با C[i] است به ازای C[i] C[i] نام مقدارشان برابر با C[i] است به ازای C[i] به ازای C[i] است که مقدارشان برابر با C[i] است به ازای C[i] است به ازای C[i] به ازای C[i] است به ازای C[i] است به ازای C[i] به ازای C[i] است به ازای C[i] به از ازای C[i] به از رابر با با ازای C[i] به ازای C[i] به ازای C[i] به از رابر با با ازای C[i] به از رابر با با ازای C[i] به از رابر با با ازای C[i] به ازای بازای باز
 - خطوط ۷ و ۸ به ازای هر $i=0,1,\cdots,k$ تعیین میکند چند عنصر ورودی کوچکتر یا مساوی i وجود داند.
 - در نهایت حلقه خطوط ۱۱ تا ۱۳ یک بار دیگر آرایه A را بررسی میکنند، اما اینبار از آخر. در این بررسی هر عنصر A[j] در مکان درست خود در آرایه B قرار میگیرد.

- اگر همه n عنصر متمایز باشند، وقتی وارد خط ۱۱ میشویم، به ازای هر A[j] مقدار C[A[j]] مکان درست A[j] در آرایه خروجی است زیرا C[A[j]] عنصر کوچکتر یا مساوی A[j] وجود دارد.
- از آنجایی که عناصر ممکن است متمایز نباشند، حلقه مقدار C[A[j]] را در هربار که مقدار A[j] را در B قرار می دهد کاهش می دهد. کاهش C[A[j]] باعث می شود عنصر قبلی A با مقدار A[j] به مکان قبل از A[j] در آرایه B برود.

 $\Theta(k)$ برای تحلیل زمانی الگوریتم مرتبسازی شمارشی، مشاهده میکنیم که حلقه خطوط ۲ و ۳ در زمان $\Theta(k)$ انجام میشود و حلقه خطوط ۴ و ۵ در زمان $\Theta(n)$ و حلقه خطوط ۷ و ۸ در زمان $\Theta(n)$ و حلقه خطوط ۱۱ تا ۱۳ در زمان $\Theta(n)$ انجام میشوند. بنابراین زمان کل مورد نیاز $\Theta(n+k)$ است.

- در عمل تنها وقتی از مرتبسازی شمارشی استفاده میکنیم که داشته باشیم k = O(n) که در این صورت زمان اجرا برابر با $\Theta(n)$ خواهد بود.

مرتبسازی شمارشی

مقايسهاي نيست.

مرتبسازی شمارشی از مرتبسازیهای مقایسه ای که کران پایین زمان اجرای آنها $\Omega(n \lg n)$ است بهتر عمل میکند. هیچ مقایسه ای بین عناصر در این الگوریتم انجان نمی شود و بنابراین این الگوریتم یک الگوریتم

- یک ویژگی مهم مرتبسازی شمارشی این است که پایدار ¹ است، بدین معنا که عناصری که مقدارشان برابر است در آرایه خروجی با همان ترتیب ورودی ظاهر میشوند. ویژگی پایداری مهم است بدین دلیل که ممکن است علاوه بر کلیدی که مرتبسازی برروی آن انجام میشود، دادههای دیگری همراه کلید وجود داشته باشند و نخواهیم ترتیب آنها تغییر کند.

¹ stable

مرتبسازي مبنايي

م مرتبسازی مبنایی 1 در مواقعی استفاده میشود که میدانیم اعداد d رقمی هستند و هریک از ارقام k مقدار متفاوت میتوانند داشته باشند.

- هریک از ارقام اعداد در مبنای ۱۰ میتوانند ۱۰ مقدار متفاوت داشته باشند و ارقام اعداد در مبنای ۲ تنها ۲ مقدار متفاوت دارند.

- این الگوریتم اعداد را به ترتیب کم ارزشترین رقم آنها (از رقم راست به چپ) مرتب میکند.

¹ radix sort

مرتبسازی ۷۰/۶۰

ساختمان داده

مرتبسازی مبنایی

- در شکل زیر یک نمونه از این مرتبسازی نشان داده شده است.

9
5
6
7
7
0
9
(

مرتبسازي مبنايي

- برای این که این مرتبسازی درست عمل کند، الگوریتم مرتبسازی ارقام باید پایدار باشد.
- تابع Radix–Sort آرایه A[1:n] را که هرکدام از عناصر آن d رقم دارند مرتب می کند. رقم اول کمارزش ترین یعنی رقم سمت راست . رقم d ام پرارزش ترین رقم یعنی رقم سمت چپ است.

Algorithm Radix Sort

function RADIX-SORT(A, n, d)

- 1: for i = 1 to d do
- 2: use a stable sort to sort array A[1:n] on digit i

مرتبسازي مبنايي

شمارشی استفاده می شود.

- اگرچه در الگوریتم مرتبسازی مبنایی الگوریتم مرتبسازی ارقام تعیین نشده است، اما معمولا از مرتبسازی

ساختمان داده Vo /84 مرتبسازي

- قضیه : به ازای n عدد d رقمی به طوری که هررقم بتواند k مقدار داشته باشد، مرتبسازی مبنایی اعداد را در زمان $\Theta(n+k)$ انجام شود.

رمان و بنابراین تعداد d رقم در زمان $\Theta(n+k)$ مرتب میشود و بنابراین تعداد $\Theta(n+k)$ مرتب می $\Theta(d(n+k))$

- اگر d ثابت باشد و k = O(n) باشد، مرتبسازی مبنایی در زمان خطی آرایه را مرتب می کند.

مرتبسازي سطلي

مرتبسازی سطلی 1 فرض می کند دادههای ورودی دارای یک توزیع احتمالی یکنواخت 2 هستند.

مرتبسازی سطلی بازه [0,1] را به n زیربازه مساوی که هرکدام یک سطل ³ نامیده می شوند تقسیم می کند و n عدد ورودی را در این سطلها توزیع می کند. از آنجایی که داده های ورودی با توزیع احتمالی یکنواخت پراکنده شده اند، تعداد اعدادی که در یک سطل قرار می گیرند زیاد نخواهد بود. برای مرتبسازی اعداد به ترتیب از زیربازه اول شروع می کنیم و اعداد درون بازه ها را مرتب می کنیم و بدین ترتیب همه اعداد مرتب می شوند.

¹ bucket sort

² uniform distribution

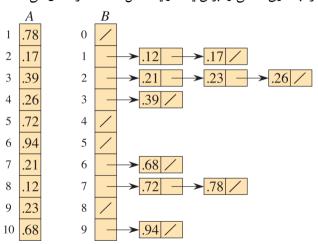
³ bucket

مرتبسازي سطلي

- تابع Bucket-Sort در زیر نشان داده شده است. این تابع آرایه A[i:n] را که هر یک از عناصر A[i] در محدوده $A[i] \gg 0$ است را دریافت می کند. این تابع از یک آرایه اضافی B[0:n-1] که سطلها هستند استفاده می کند.

مرتبسازي سطلي

- شکل زیر عملیات مرتبسازی سطلی را برای یک آرایه شامل ۱۰ عنصر نشان می دهد.



- جرای اثبات درستی این الگوریتم، دو عنصر A[i] و A[i] را در نظر بگیرید. بدون از دست دادن عمومیت اثبات، فرض کنید $A[i] \otimes A[i] \otimes A[i] \otimes A[i]$ ، یا عنصر $A[i] \otimes A[i] \otimes A[i]$ اثبات، فرض کنید و یا به یک سطل با اندیس کمتر وارد می شود. اگر A[i] و A[i] به یک سطل یکسان بروند، حلقه خطوط ۶ و ۷ ترتیب درست را به دست می آورد و اگر در دو سطل متفاوت وارد شوند، خط ۸ آنها را در ترتیب درست قرار می دهد. پس الگوریتم درست عمل می کند.
 - برای تحلیل زمان اجرای این الگوریتم، مشاهده کنید که همه خطوط به غیر از خط V در زمان O(n) اجرا می شوند. بنابراین باید زمان اجرای n فراخوانی تابع مرتبسازی درجی را در خط V را بیابیم.

- فرض کنید n_i یک متغیر تصادفی باشد که تعداد عناصر در سطل B[i] را نشان میدهد. از آنجایی که مرتبسازی درجی در زمان مربعی اجرا میشود، زمان اجرای مرتبسازی سطلی برابر است با

$$\mathsf{T}(\mathsf{n}) = \Theta(\mathsf{n}) + \sum_{i=0}^{\mathsf{n}-1} \mathsf{O}(\mathsf{n}_{\mathsf{i}}^2)$$

 $\mathsf{T}(\mathsf{n}) = \Theta(\mathsf{n})$ با محاسبه این عبارت به دست می آوریم