HW2_Report 109550027 紀竺均

Part one: sortIntroduction

Introduction of different method/ algorithm

這部分實作了兩個 sort,Quicksort 和 Mergesort,兩者都是 Divide and conquer algorithm。Quicksort 每一輪將最右邊的點作為 pivot,反覆將 array 利用 pivot 分為左右兩邊,最後整理好整個 array;而 Mergesort 則是將 array 不斷分成兩半直到最小的單位,接著把重複把兩邊 Merge 在一起,相對於 Quicksort,Mergesort 需要創造額外的空間給被分開的 array。

- Implement Details
- o Steps

Quicksort

主要有兩個 function: quicksort()和 partition(),partition()的功能就是利用交換將 array 分成 pivot 的左右兩邊,最後 return 一個 i 值紀錄 pivot's index,而 quicksort()一直不斷遞迴的partition 並將 array 分成 i 的兩左右邊,最後重新進行quicksort。

Mergesort

主要有兩個 function: mergesort()和 merge(), mergesort()裡面包含分割的概念,直到不能繼續分割後才會由小到大開始merge; merge()先創造兩個 array 將左右兩邊傳入,最後進行比較大小。

比較大小是如果 left 比較小,就加回 A 裡面,i++,反之亦 同。

Time.h 的部分,用 clock_t starttime = clock()以及 clock_t endtime = clock()去包住我要測量的時間,然後再用 float(endtime-starttime) / CLOCKS_PER_SEC(=1000)就可以得到時間差的秒數。

- Results
 - Quicksort 測資數:5 時間:0



測資數:516 時間:0.002s

CM: Chienry Chuch Desktoph WPZ_109550027 Quicksort_109550027 Qu

Ⅰ Mergesort 測資數:5 時間:0

```
■ CAUsers\chuch\Desktop\HW2_109550027\Mergesort_109550027.exe

5
7 - 1 - 2 4 9
- 2 - 1 4 7 9
spent time: 0 seconds
7
8 4 5 - 10 3 4 - 6
- 10 - 6 3 4 4 5 8
spent time: 0 seconds
0

Process exited after 1.043 seconds with return value 0
請按任意鍵繼續 . . . ■
```

測資數:516 時間:0

```
$16
906. 9912. 9810. 9808. 9782. 9776. 9771. 9726. 9714. 9709. 9695. 9683. 9645. 9639. 9590. 9557. 9510. 9490. 9484. 9447. 9404. 9321. 9312. 9310. 9309. 9283. 9266. 9264. 9164. 9135. 9108. 9076. 9073. 9029. 8970. 8938. 8891. 8885. 8836. 8764. 8731. 8672. 8666. 8614. 8586. 8583. 8572. 8560. 8426. 8366. 8356. 8356. 8335. 8322. 81815. 8105. 8102. 8085. 8028. 8026. 8010. 7864. 7852. 7843. 7822. 7818. 7755. 7724. 7720. 7708. 7647. 7610. 7534. 7519. 7431. 7436. 7413. 7535. 7734. 7720. 7708. 7647. 7610. 7534. 7519. 7436. 7418. 7512. 7520. 7520. 7520. 7520. 8426. 8426. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 8436. 84
```

測資書:652 時間:0.001s

- Time complexity
 - Quicksort: O(nlogn)
 - Mergesort: O(nlogn)
- Discussion
- o Your discover

Sorting 的部分好像蠻容易的,但是覺得他蠻厲害的畢竟用 bubble sort 什麼的很笨。然後計算 time 的部分真的快搞死我,他一開始時間差全部都是零,我自我懷疑超久,後來尋求助教協助,助教說是我的測資不夠大。之後跑很大的測資才開始有一些數字跑出來,可是我要測卻沒辦法一次貼上那麼多,例如測資二測資三,只能一次貼上一筆數據。

Which is better algorithm in which condition

如果相同的測資比較多,或者是 reverse case,用 quick sort 比較快。空間不夠的話,用 quick sort 也比較省空間。

- o Challenges you encountered
 - Mergesort

- 一開始沒加藍色框框,在其中一邊已經全部加進去,另一邊還沒加完的情況下,另一邊沒辦法加進去,結果出現很多 0。
- Conclusion
- O What did you learn from this homework

- O How many time did you spend on this homework
- Feedback to TA

大概都花了兩三個小時,學到最重要的就是這兩個演算法,對於降

低 time complexity 很有幫助。也學到了 time.h 的應用。

Part two: MST

• Introduction

Introduction of different method/ algorithm

這部分實作了兩個 MST: Prim 和 Kruskal,兩者都 Greedy algorithm,

Prim 著重在 vertices 上;Kruskal 著重在 edges 上。

- Implement Details
- o Steps
 - Prim

創建了一個 class: graph,把 prim 的所有功能包在裡面。Graph有兩個變數 v for vertex index, vector<intpair> array myadj。 myadj[]裡面用一個 intpair(int,int)存取 adjacent vertex 的 index 和自己與該 vertex 之間的 edge cost。

利用 addedge()將 endnode push back 到 startnode 的 myadj,

同時也將 startnode push_back 到 endnode 的 myadj。在 prim()裡面,先建立一個 minheap pririty queue,存取 intpair (key, vertex_index),key 存在前面是因為 priority queue 會依照第一個 element 的大小去 sort queue。接著建立三個 vector: <book> isvisit, <int> parent, <int> key,並讓 key[0]=0,其他= infinity,之後 push 到 priority queue 內。

最後透過 while 迴圈以及迭代器,將 pop 掉的 unvisit vertex 的 adgacent vertex 進行更新 key 和 parent 的動作,更新依據: key[鄰居]是否大於 edge weight,最後將更新後的 key push 回

priority queue,以便下回可以再將 weight 最小的 unvisit vertex pop 掉。

Kruskal

首先,我利用創了一個 std::set<int>的 array:sets[],每一個 vertex 都是一個 set。

再來,我創了一個 class: edge,裡面存放三個 int 變數:
startnode, endnode, weight,以及一個 function:sort edge(),
為了比較 edge 的大小,我 overloading "<" operator,告訴他
edge 的大小想要依照 weight 的大小做比較,並使用 std::sort
去處理。

接著,我寫了一個 int function: find set()。傳入的值有 sets array, sets array size, 和我要找的 vertex v。利用 for 迴圈以及 find(),若我要找的 v 在 sets[i]裡面,就會傳回 i 的值。接著,在 main 裡面用 for 迴圈讓每一個 edge 跑一次(weight 從低至高的 edge),讓每個 edge 的 startnode 及 endnode 做 find set,若兩者 sets[i]的 i 值不同,則進行 union 的動作。 Union function:若兩個 set 不相同,則將比較小的 set 併入比較大的 set。我利用 set1.insert(set2.begin(),set2.end())將其合併,還要記得 set2.clear()才不會在下次 find set 出現問題。

Results

■ Prim(測資一) (vertex,edge)=(1000,99822)

結果:58630 Time complexity:O(E logV)

```
■ C\Users\chuch\Desktop\\HW2_109550027\Prim_109550027.exe

999 681 9656
999 858 7357
999 920 1955
999 186 7552
999 545 6361
999 596 5234
999 990 83 630
999 990 1882
999 430 459
999 613 4780
999 137 5354
999 402 6267
999 65 3219
999 863 1334
999 402 6267
999 65 3219
999 758 4106
999 389 5911
999 778 5616
999 379 5812
999 529 5326
999 178 5616
999 47 9454
999 639 5280
999 649 5537
999 226 9585
999 799 799 1321
999 318 1843
999 33 5439

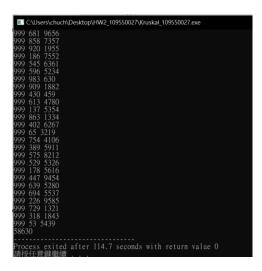
Process exited after 113.5 seconds with return value 0

□特安任意鍵繼續 . . .
```

Kruskal

(測資一) (vertex,edge)=(1000,99822)

結果: 58630 Time complexity: O(E logE)



Discussion

o Your discover

最重要的發現是,很多東西我原本以為他只是一個名字,要自己去 定義,結果查了 cplusplus.com 以後才發現他們都是內建的 function... 舉例 priority_queue、set 等等,她們就像好工具一樣會讓速度快很多,工欲善其事必先利其器,所以平常還是要多打扣來增廣見聞。
再來就是我發現除了之前交的 sorting 之外,還有很多 function 都附帶有排序的功能,例如上面講的 priority_queue 和 set,除此之外還發現 std::sort 不一定只能按數字大小排,也可以透過 function 或者是 object 來排(第三個 parameter),而我這次選擇了 operator overloading,就不需要第三個 parameter。
o Which is better algorithm in which condition當 V>>E 時,kruskal的 Time complexity較小,反之則 Prim較小。o Challenges you encountered

Prim

原本我一直 push 新的 key 值到 priority queue,但舊的還沒有 pop 掉(也不知道怎麼 pop 因為他不是最小的值,例如 infinity) 後來就發現可以先讓 current v = pop 掉的值,再接 if current v isvisit: continue,這樣子那些 inf 或其他值最終還是會 pop 完!

Kruskal

在寫 union()時,一直無法對我的 set 做改變,後來索性把它丟到 main 裡面然後就好了,後來想想應該是沒有用好 pointer 的關係。還有一個讓我找超級久的 bug,就是在確認 findset(start)!=findset(end)之後,一開始我讓 a=edge.start, b=edge.end,然後去比較還有 union sets[a], sets[b],但後來我

發現 a 和 b 應該要是 find set 之後的值,而不是 edge 的值,真的好笨喔而且他好難找……

- Conclusion
- What did you learn from this homework
- O How many time did you spend on this homework

跑而且跑出對的就覺得好開心,但我還是很累。

o Feedback to TA

■ Prim

大概花五六個小時左右,學到了什麼是 pair, make pair, 還有 怎麼樣定義整數的無限大(但不是真的無限大),也學了 priority queue 和怎麼把它改成小到大("greater),除此之外,本來對 iterator 沒有很了解,也因為這次機會去了解了。

Kruskal

花了七八個小時吧,雖然題目也沒有到很難主要是 debug de 超久。學到的東西: sort, set 一些小東西(都打在 discover 了)。 結論:這兩題都打好久喔連 report 也是,但是看到那一大串測資在 Part three: shortest path

- Introduction
- Introduction of different method/ algorithm

這部分實作 single source shortest path,和 MST 的差別是不同的起始 點會造成不同結果。

- Implement Details
- o Steps
 - Dijkstra

因為 dijkstra 也是著重在 vertices 上,所以一樣使用 Class Graph 以及兩個變數 v for vertex index, vector<intpair> array myadj。 myadj[]裡面用一個 intpair(int,int)存取 adjacent vertex 的 index 和自己與該 vertex 之間的 edge cost。

利用 addedge()將 endnode push_back 到 startnode 的 myadj,因為是 directed weight 所以不必將 startnode push_back 到 endnode 裡。

在 dijkstra()裡面,先建立一個 minheap pririty queue,存取 intpair (distant, vertex_index)。接著建立三個 vector: <book isvisit, <int> parent, <int> distant, 並讓 starting vertex 的 distant=0, 其他= infinity,之後 push 到 priority queue 內。 利用 while 迴圈,若 Q!=empty(),也就是說若還沒 visit 完全部 vertex,則依序將 vertex pop()並且更新他的鄰居們。更新的依據:如果 distant[current vertex]+edge weight<distant[鄰居],就

讓鄰居的 distant = distant[current vertex]+edge weight。 最後將 current vertex 的 isvisit 改為 1。最後因為題目要求印出 s 到 t 的 shortest path,其實就是印出 t 的 distant 值。

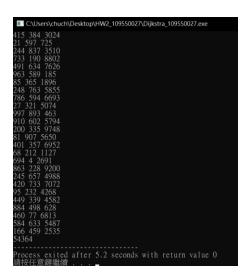
■ BellmanFord

Bellmanford 是從 edge 的角度去更新資料,所以沿用 class: edge,裡面存放三個 int 變數:startnode, endnode, weight。 在 bellmanoford()裡面,建立一個 vector distant 存取 vertex 的 資料,讓 starting vertex=0,其他= infinity。 接著進行 V-1 次的資料更新。對所有 edge,檢查 distant[start] + weight 是否小於 distant[end],若較小,則使 distant[end] = distant[start] + weight。 (negative-weight cycle detect 在 discussion)。最後,只要印出 t 的 distant 值就好了。

Results

■ Dijkstra(測資二) (vertex,edge)=(1000,2000)

結果:54364 Time complexity:O(V^2)



BellmanFord

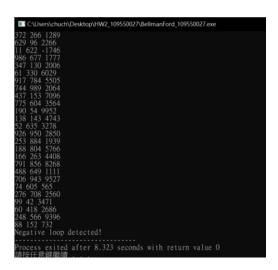
(測資二) (vertex,edge)=(1000,2513)

結果: 24585 Time complexity: O(VE)



(測資三) (vertex,edge)=(1000,3000)

結果: Negative loop detected! Time complexity: O(VE)



Discussion

o how to print out the path of the shortest path

這是一個很好的問題,因為我本來有寫一個 vector 去存 parent,可是後來發現都沒有用到就把他刪了。回到問題,如果我要印出整個 path 的話,我必須要知道她的 parent 是誰,然後從兒子一直往

parent 推,直到推回 starting vertex。

- o Your discover
 - BellmanFord
 - How to detect negative loop

結束 V-1 次的檢查後,最後再進行第 V 次的檢查,正常情況下 V-1 次就可以產生 shortest path,如果第 V 次數據還會改變的 話,也就是 distant[start] + weight < distant[end],就代表有 negative-weight cycle,才會讓他每跑一次都比本來值的小。

oTime complexity

- Dijkstra: O(V^2)
- BellmanFord: O(VE)
- Which is better algorithm in which condition

如果 V 比 E 大很多的話,用 bellman ford 會比較快;此外,如果有 negative weight cycle 的可能的話,一定不能選擇 dijkstra。(bellman ford 也只能偵測它的存在而已,不能計算。)

- o Challenges you encountered
 - BellmanFord:一開始 if 裡面沒有寫 if distant[start]!=inf,然後結果一直不對,把過程印出來以後發現他的數字都沒改,猜想可能是在最大 int_max 附近做加減導致問題。還有本來是照著講義,negative weight cycle 的判斷式 if distant(end) + weight != distant(start) ,結果怎麼跑結果都是 negative weight cycle,後來想到應該要改成小於才對,因為本來就不會每個點都是shortest path 更不可能永遠左右相等。

- Conclusion
- What did you learn from this homework
- O How many time did you spend on this homework
- o Feedback to TA

心得:寫的時候,Dijkstra 是參考 prim,因為兩者都是用 vertex 做比較;而 bellman 是參考 kruskal,他們都是讓 edge 跑迴圈,比較 start node 及 end node。因為先寫了 MST,讓這兩題變得容易許多,大概都花一個小時就完成了,其中最困難的應該就是怎麼拼出 Dijkstra 了吧!