The Greedy Method

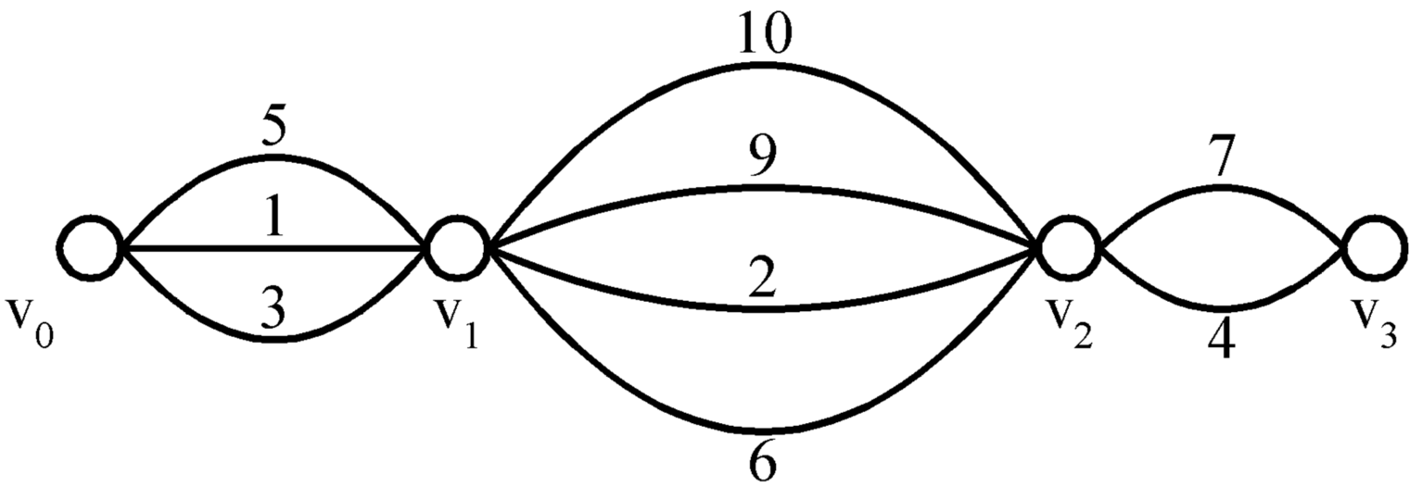
貪婪法

The Greedy Method

- 解最佳化問題的演算法，其解題過程可看成是由一連串的決策步驟所組成，而每一步驟都有一組選擇要選定。
- 一個 **Greedy method** 在每一決策步驟總是選定那目前看來最好的選擇。
- **Greedy methods** 並不保證總是得到最佳解，但在有些問題卻可以得到最佳解。

The Greedy Method

- E.g. Find a shortest path from v_0 to v_3.
  - The greedy method can solve this problem.
  - The shortest path: 1 + 2 + 4 = 7.



Part 1

Outline

Greedy Method

- Greedy Algorithm for Vertex-coloring
- ST (Spanning Tree) Algorithms
  - BFS algorithm
  - DFS algorithm
- MST (Minimum Spanning Tree) Algorithms
  - Kruskal's algorithm
  - Prim's algorithm
- Shortest Path (Distance Tree) Algorithms
  - Dijkstra's algorithm

The Greedy Algorithm for Vertex-Coloring

Let G be a graph in which the vertices have been listed in some order $x_1, x_2, \ldots, x_n$.

(1) Assign the color 1 to vertex x_1.

(2) For each $i = 2, 3, \ldots, n,$ let p be the smallest color such that none of the vertices $x_1, x_2, \ldots, x_{i-1}$ which are adjacent to x_i is colored p, and assign the color p to x_i.

Brooks' theorem,1941

- **If G is a connected graph which is neither a complete graph nor an odd cycle, then $\chi(G) \leqq \Delta(G)$.**

($\Delta(G)$ is the maximum degree of a vertex in G)

- The Greedy algorithm for vertex-coloring derive that $\chi(G) \leqq \Delta(G)+1$.

Algorithm for Spanning Tree rooted at u

Let G=(V,E) be a graph of order n and let u be any vertex.

(1) Put U=$\{u\}$ and F=$\varnothing$.

(2) While there exists a vertex x in U and a vertex y not in U such that $\alpha=\{x, y\}$ is an edge in G.

(i) Put the vertex y into U.

(ii) Put the edge α into U.

(3) Put T=(U,F).

Breadth-First Algorithm for Spanning Tree rooted at u

Let G=(V,E) be a graph of order n and let u be any vertex.

(1) Put $i=1$, U=$\{u\}$, D(u)=0, $bf(u)=1$, F=$\varnothing$, and T=(U,F).

(2) If there is no edge in G that joins a vertex x in U to a vertex y not in U, then STOP!

Otherwise, determine an edge $\alpha=\{x,y\}$ with x in U and y not in U such that x has the smallest BF number $bf(x)$, and do the following:

(i) Put $bf(y) = i+1$.

(ii) Put D(y) = D(x) + 1.

(iii) Put the vertex y into U.

(iv) Put the edge α into U.

(v) Put T=(U,F).

(vi) Increase i by 1 and go back to (2)

Depth-First Algorithm for Spanning Tree rooted at u

Let G=(V,E) be a graph of order n and let u be any vertex.

(1) Put i=1, U={u}, df(u)=1, F=∅, and T=(U,F).

(2) If there is no edge in G that joins a vertex x in U to a vertex y not in U, then STOP!

Otherwise, determine an edge $\alpha=\{x, y\}$ with x in U and y not in U such that x has the largest DF number $df(x)$, and do the following:

(i) Put $df(y) = i+1$.

(ii) Put the vertex y into U.

(iii) Put the edge α into U.

(iv) Put T=(U,F).

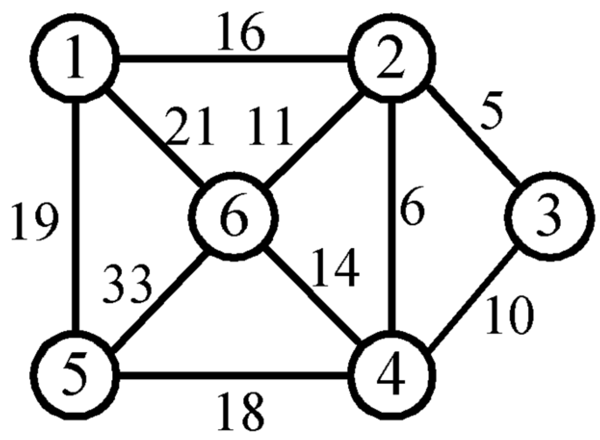
(v) Increase i by 1 and go back to (2)

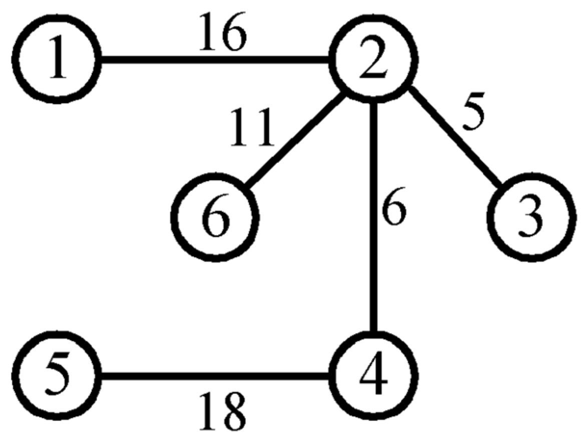
Minimal Spanning Trees

- It may be defined on Euclidean space points or on a graph.
- $G = (V, E)$: weighted connected undirected graph
- **Spanning tree** : $S = (V, T)$, $T \subseteq E$, S is a undirected tree
- **Minimal spanning tree (MST)** : a spanning tree with the smallest total weight.

Minimal Spanning Trees

- A graph and one of its minimum spanning trees (MSTs)





圖的最小生成樹(MST)

- 一個最小生成樹演算法 — **Kruskal**演算法，採用貪婪法(greedy method)的策略來建構最小生成樹，也就是每次都是挑選最小成本且不形成cycle的邊加入最小生成樹T之中，如此經過n-1次的邊的挑選之後形成的累積成本必定是最小
- 另一個採用貪婪法(greedy method)的策略稱為**Prim’s**演算法，也就是每次都是挑選最小成本的邊加入最小生成樹T之中，經過n-1次的挑選之後形成的累積成本必定是最小。由於每次挑選邊時，都是挑選一個具有連結X及V-X的邊，也就是說挑一個一頂點在X，而另一個頂點在V-X的邊，因此，將所挑選的邊加入T之後不會形成cycle，這代表T是一棵樹(tree)。

Kruskal's Algorithm for Minimum Spanning Tree

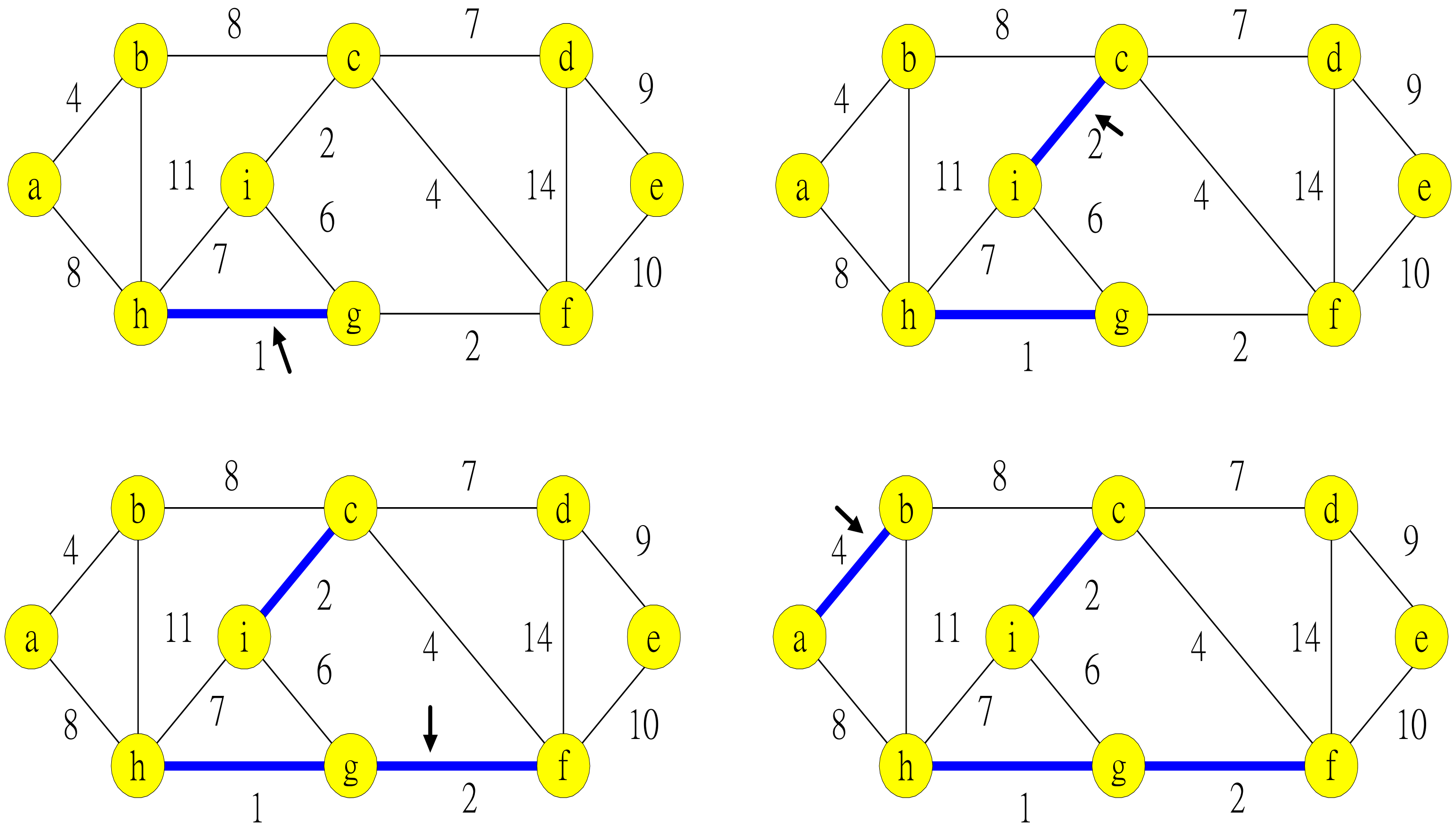
Let G=(V,E) be a weighted connected graph with weighted function *c*.

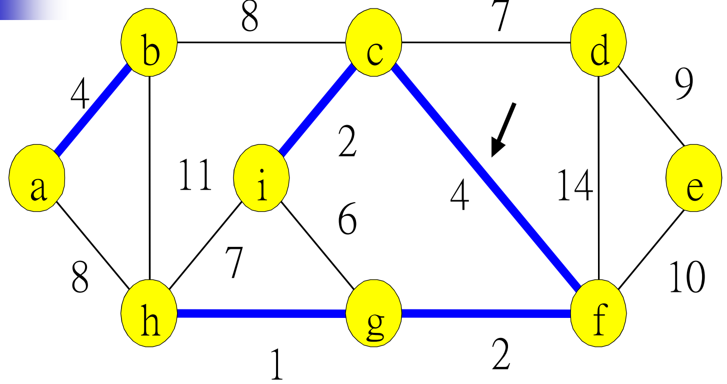
(1) Put F=$\varnothing$.

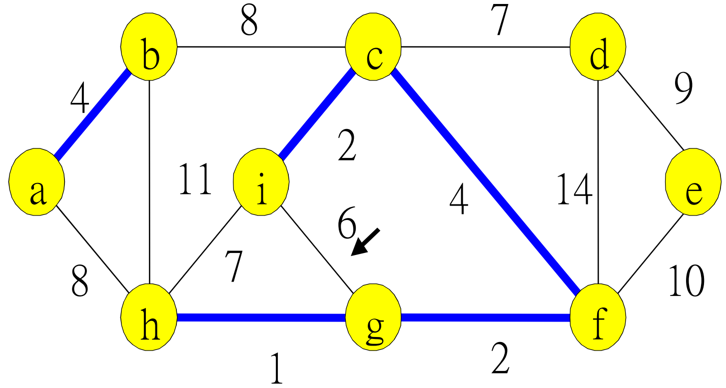
(2) While there exists a edge α not in F such that F $\cup$ {α} do not contain the edges of a cycle of G, determining such an edge α of minimum weight and put α in F.

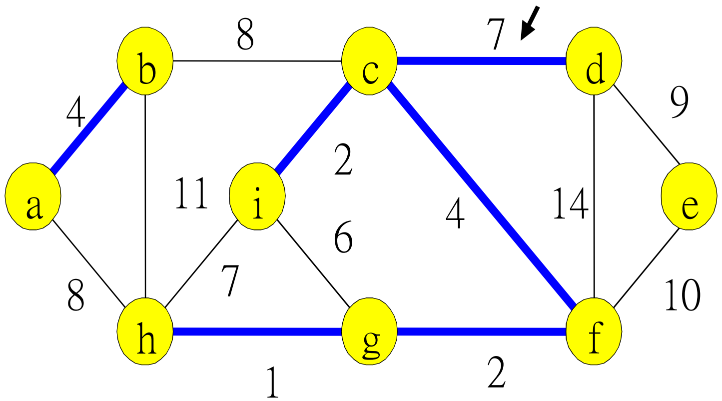
(3) Put T=(V,F).

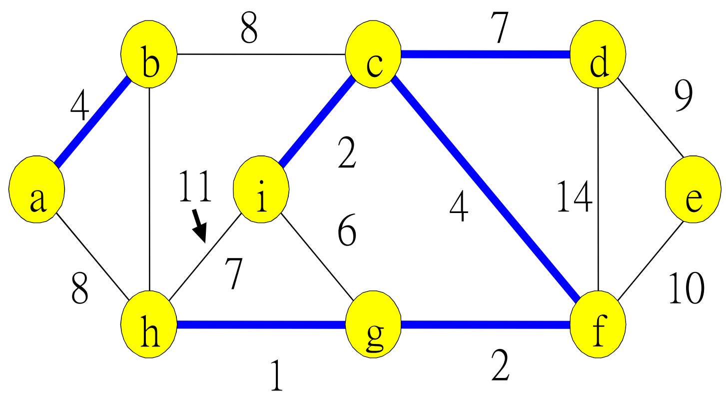
Kruskal's Algorithm -Construct MST

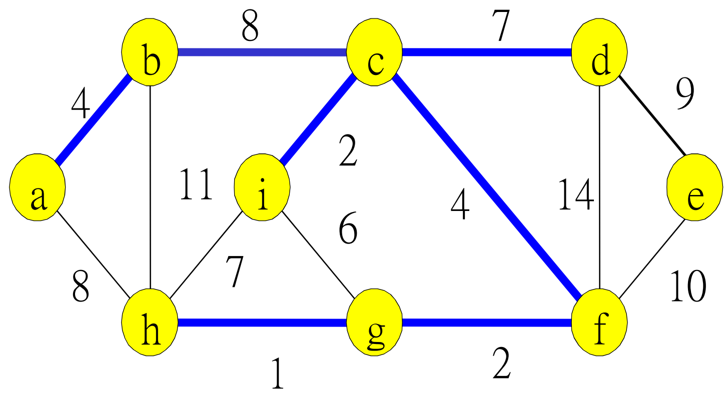


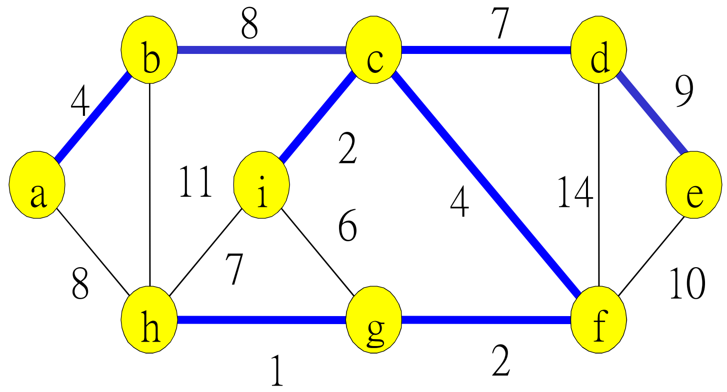

a
b
c
d
e
f
g
h
i
8
7
4
9
11
2
4
14
6
8
7
10
1
2


a
b
c
d
e
f
g
h
i
8
7
4
9
11
2
4
14
6
8
7
10
1
2


a
b
c
d
e
f
g
h
i
8
7
4
9
11
2
4
14
6
8
7
10
1
2


a
b
c
d
e
f
g
h
i
8
7
4
9
11
2
4
14
6
8
7
10
1
2


a
b
c
d
e
f
g
h
i
4
8
7
9
11
2
4
14
6
7
8
10
1
2


a
b
c
d
e
f
g
h
i
4
8
7
9
11
2
4
14
6
7
8
10
1
2

Prim's Algorithm for Minimum Spanning Tree

Let $G=(V,E)$ be a weighted connected graph with weighted function c and let u be any vertex of G.

(1) Put $i=1$, $U_1=\{u\}$, $F_1=\varnothing$, and $T_1=(U_1,F_1)$.

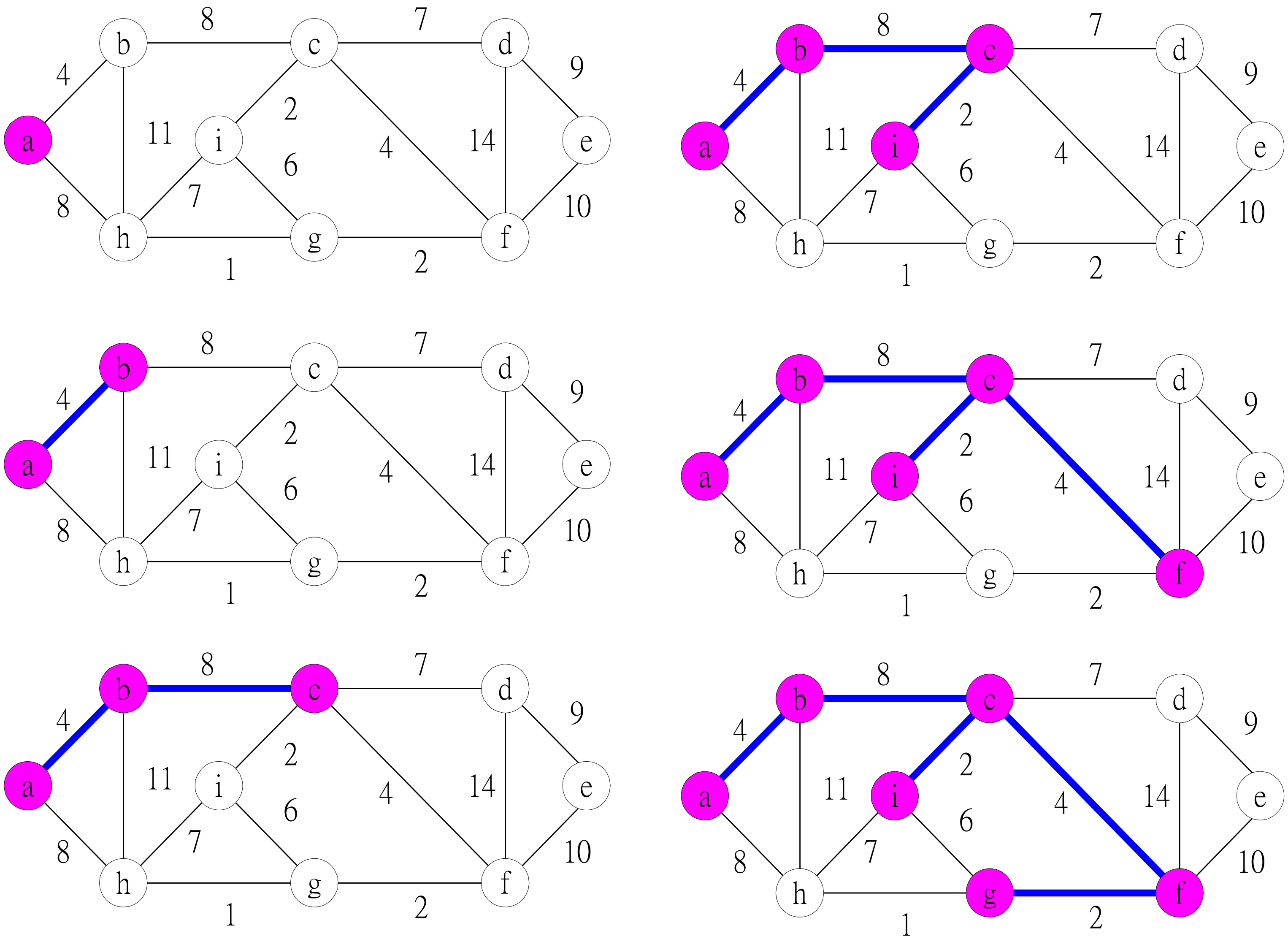
(2) For $i=1,2,\ldots,n-1$, do the following:

(i) Locate an edge $\alpha_i=\{x,y\}$ of smallest weght such that x is in U_i and y is not in U_i.

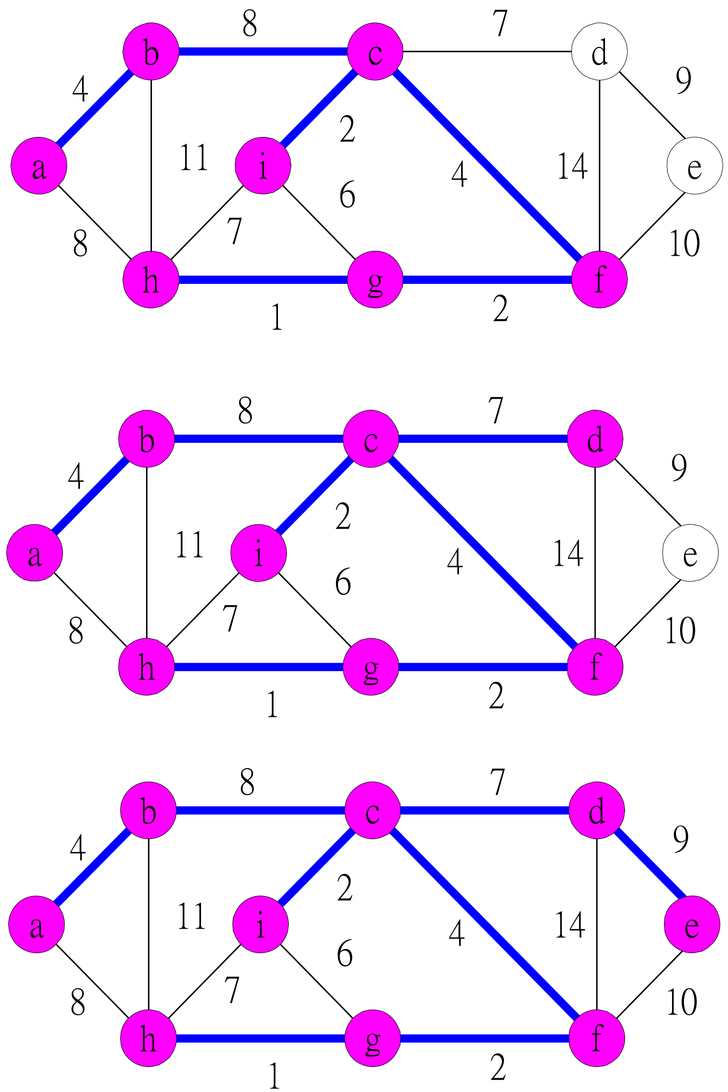
(ii) Put $U_{i+1}=U_i\cup\{y\}$, $F_{i+1}=F_i\cup\{\alpha_i\}$ and $T_{i+1}=(U_{i+1},F_{i+1})$.

(iii) Increase i to $i+1$.

(3) Output $T_{n-1}=(U_{n-1},F_{n-1})$. (Here $U_{n-1}=V$.)



b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2
b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2
b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2
b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2
b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2
b 8 c 7 d 4 9 a 11 i 2 4 14 e 6 8 7 10 h g f 1 2


a
b
c
d
e
f
g
h
i
4
8
7
9
11
2
14
7
6
4
10
8
1
2

圖的最短路徑

- 由圖中的某個頂點(vertex)v到圖中的另一頂點u，若v到u之間存在一條路徑(path)，則路徑中所經過的邊(edge)的加權值(weight)的總合稱為路徑的成本(cost)。所有路徑中具有最小成本的稱為最短路徑(shortest path)。

- 由於最短路徑具有許多應用，因此有許多求取最短路徑的演算法，我們介紹最有名的演算法：Dijkstra演算法。

Dijkstra's Algorithm for a Distance Tree rooted at u

Let G=(V,E) be a weighted graph of order n with weighted function c, and let u be any vertex.

(1) Put $U=\{u\}$, $D(u)=0$, $F=\varnothing$, and $T=(U,F)$.

(2) If there is no edge in G that joins a vertex x in U to a vertex y not in U, then STOP!

Otherwise, determine an edge $\alpha=\{x,y\}$ with x in U and y not in U such that $D(x)+c\{x,y\}$ is as small as possible, and do the following:

(i) Put the vertex y into U.

(ii) Put the edge α into U.

(iii) Put $D(y)=D(x)+c\{x,y\}$ and go back to (2)

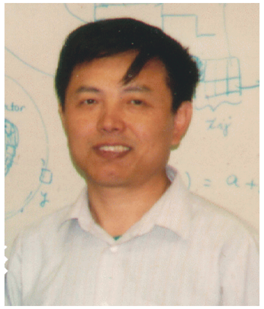
要證明演算法是找到最佳的結果!

Part 2

Steiner Ratio Conjecture

Gilbert and Pollack,1968

Steiner Ratio Conjecture is True



- Ding-Zhu Du (堵丁柱) and Frank Kwang-Ming Hwang (黄光明). A Proof of the Gilbert-Pollak Conjecture on the Steiner Ratio. *Algorithmica* 7(2–3): pages 121–135, 1992. (received April 20, 1990.)

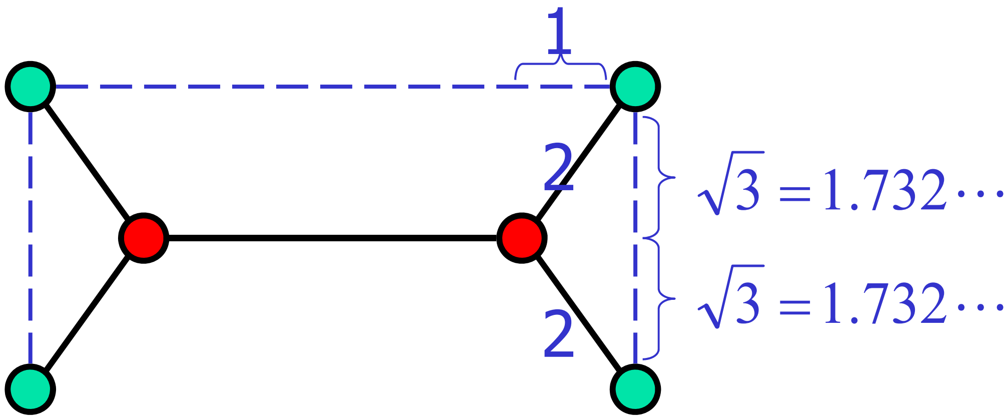
Euclidean Steiner Problem

- Given: a set *P* of <u>*n* points</u> in the <u>euclidean plane</u>
- Output: a shortest tree connecting all given points in the plane

- The tree is called a Steiner minimal tree.
- Notice that Steiner minimal trees may have extra points (Steiner points).

Minimum Spanning Tree (MST)

Steiner Minimum Tree (SMT)



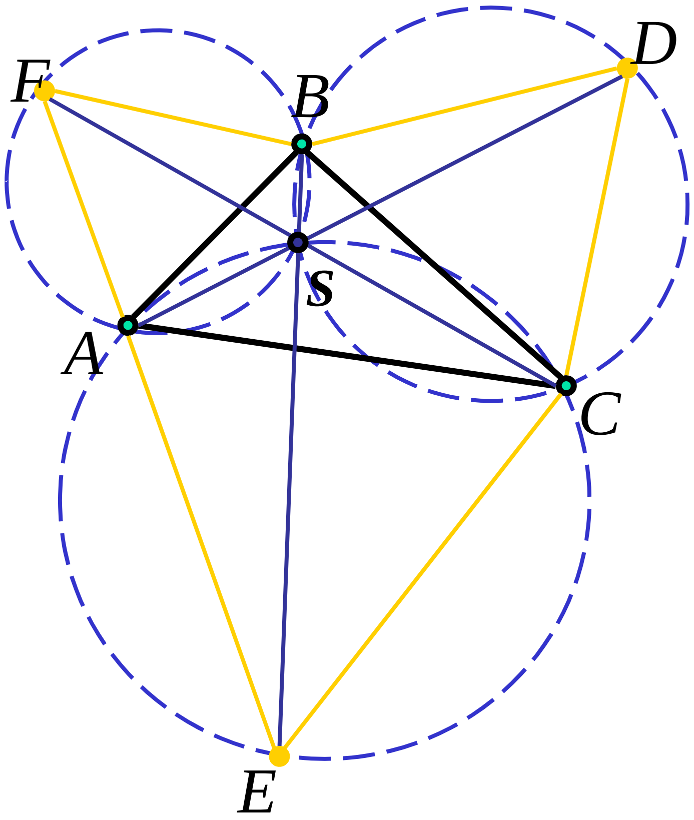
- Regular points (●): P
- Steiner points (●): $V(SMT(P)) - P$

History of SMT

- Fermat (1601-1665): Given three points in the plane, find a fourth point such that the sum of its three distances to the three given points is minimal.
- Torricelli solved this problem before 1640.

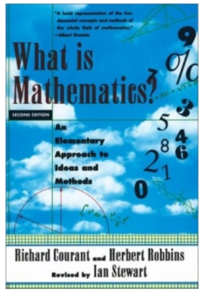
History of SMT (cont.)

- Torricelli Point (or called (First) Fermat Point):



History of SMT (cont.)

- Jarník and Kössler (1934) formulated the following problem: *Determine the shortest tree which connects given points in the plane.*
- Courant and Robbins (1941) described this problem in their classical book "*What is Mathematics?*" and contributed this problem to Jakob Steiner, a mathema at the University of Berlin in the 19th century.



The Complexity of Computing Steiner Minimum Trees

- NP-Hard!
  - M.R. Garey, R.L. Graham, and D.S. Johnson. The complexity of computing Steiner minimum trees. *SIAM J. Appl. Math.*, 32(4), pages 835–859, 1977.

➢ Approximation

Steiner Ratio

- $L_s(P)$: length of Steiner Minimum Tree on P
- $L_m(P)$: length of Minimum Spanning Tree on P
- Steiner ratio ρ = $\inf_P \left\{ \dfrac{L_s(P)}{L_m(P)} \right\} = \dfrac{\sqrt{3}}{2} \approx 0.866$

Gilbert-Pollak Conjecture

- Gilbert and Pollak conjectured that for any *P*.

$$L_s(P) \geq \frac{\sqrt{3}}{2} L_m(P) \qquad \left(\frac{\sqrt{3}}{2} \approx 0.866 \right)$$

  - E.N. Gilbert and H.O. Pollak. Steiner minimal trees. *SIAM J. Appl. Math.,* Vol. 16, pages 1–29, 1968.

Previous Results for n = 3

- n = 3: Gilbert and Pollak (1968)

Let S be the Torricelli point of $\triangle ABC$.
Suppose that $|AS| = \min\{|AS|,|BS|,|CS|\}$.

$$L_s(A,B,C) = AS + BS + CS$$

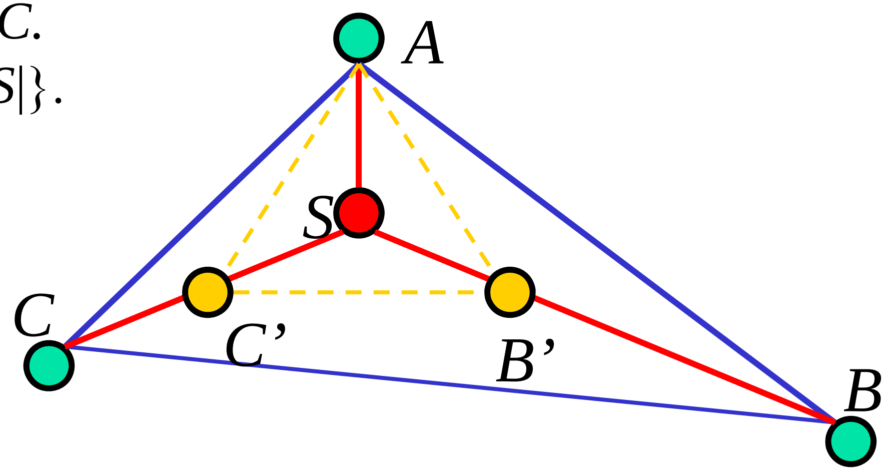
$$= AS + B'S + C'S + BB' + CC'$$

$$= L_s(A,B',C') + BB' + CC'$$

$$= \frac{\sqrt{3}}{2}(AB'+AC') + BB' + CC'$$

$$\geq \frac{\sqrt{3}}{2}(AB'+BB'+AC'+CC')$$

$$\geq \frac{\sqrt{3}}{2}(AB+AC) = \frac{\sqrt{3}}{2}L_m(A,B,C)$$

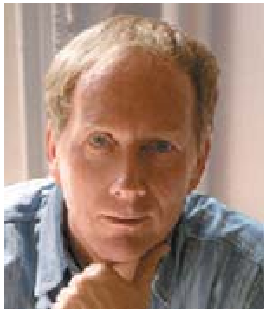


Previous Results for small *n*

- $n = 3$: Gilbert and Pollak (1968)
- $n = 4$: Pollak (1978)
- $n = 5$: Du, Hwang, and Yao (1985)
- $n = 6$: Rubinstein and Thomas (1991)

Previous Results for lower bound of ρ

- ρ = 0.5 : Gilbert and Pollak (1968)
- ρ = 0.577 : Graham and Hwang (1976)
- ρ = 0.743 : Chung and Hwang (1978)
- ρ = 0.8 : Du and Hwang (1983)
- ρ = 0.824 : Chung and Graham (1985)



Gilbert-Pollak conjecture is true!

D.-Z. Du and F.K. Hwang: Gilbert-Pollak conjecture on Steiner ratio is true, *Proceedings of National Academy of Sciences U.S.A.*, 87 (1990) 9464-9466. (Also in *Proceedings of 31st FOCS*, 1990, pp76-85 and in *Algorithmica* 7 (1992) 121-135.)

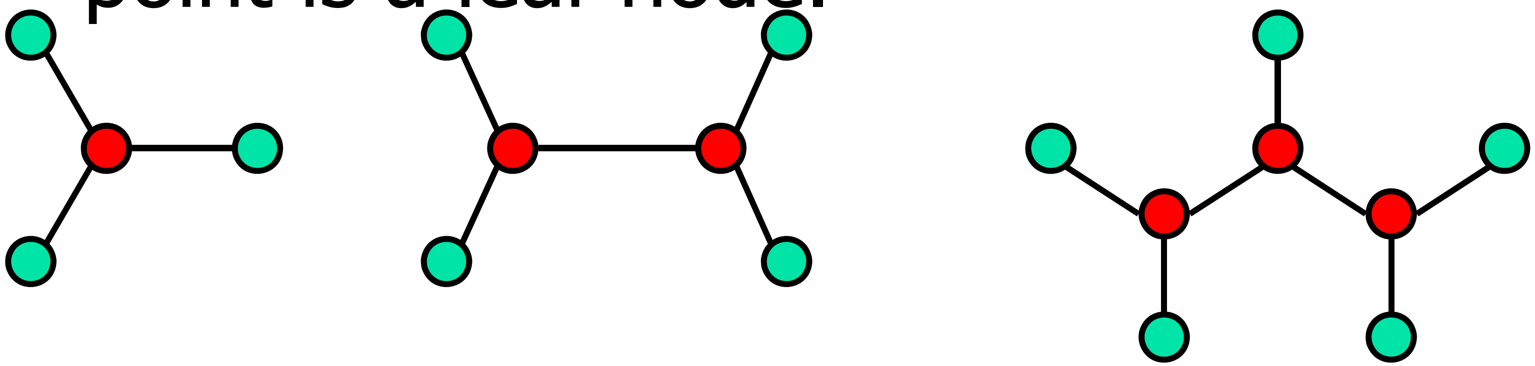
1992 The proof of Gilbet-Pollak conjecture was selected by 1992 Year Book of Encyclopaedia, Britannica, as the first one among six outstanding achievements in mathematics in 1991.

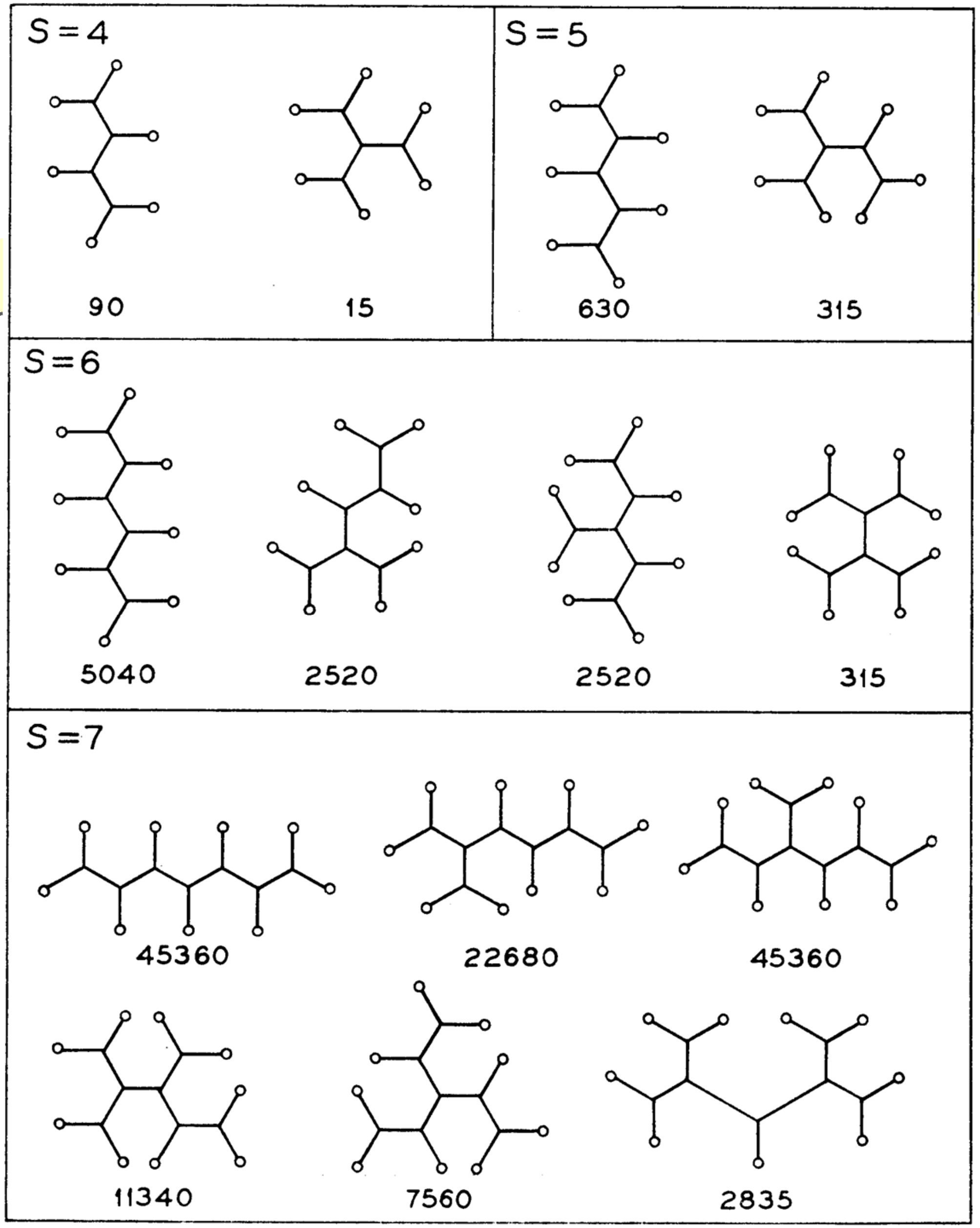
1992 Received $500 personal award from Professor Ronald L. Graham, the President of American Mathematics Society, for proving the Steiner ratio conjecture of Gilbert and Pollak.

1990-1991 The proof of Gilbet-Pollak conjecture was reported in *New York Time* on 10/30/1990, *Science* (1990, pp.1081-1082) , *Science News* (12/22-29/1990, pp.389), *SIAM News* Vol 24 No 1 (1991), *New Scientists* (April 1991, pp.22), and *New Scientists* (November 1991, pp.26-29).

Full Steiner Tree

- A Steiner tree with *n* regular points is called a full Steiner tree if it has exactly *n*-2 Steiner points, i.e., every regular point is a leaf node.




S=4
90
15
S=5
630
315
S=6
5040
2520
2520
315
S=7
45360
22680
45360
11340
7560
2835

附錄

旗落帆升

旗落帆升

●黃光明

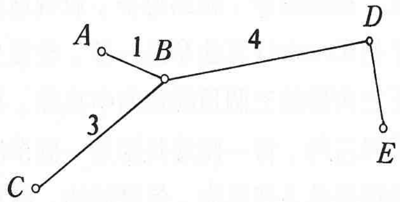
常常有人問我，你一個搞數學的人，爲什麼在電話公司一呆二十幾年？我說，爲的是在北卡州一片野地裏的那一根旗杆。什麼？旗杆？你在開玩笑！不是玩笑，朋友，請聽我說。

＊＊　　　　＊＊

一個大行號在全國各地有分店、倉庫、辦公室…等等，需要一個通訊網絡把這些據點連接起來。電話公司把這樣一個網絡叫做「私家網絡」。雖然電話公司並不眞正爲一個私家網絡去舖設線路，而仍是用公共網絡的一些線路；但是從收費的立場而言，我們可以當作一個私家網絡的線路都是爲這行號特設的，收費的多寡也即取決於這些線路的總長。因此無論從行號節省開支或電話公司公平收費的觀點，我們都需要知道如何造一個最短的網絡來連接一個行號的所有據點。這個最短網絡通常稱爲「最短生成樹」。

最短生成樹的造法其實早在一九二六年就被一個捷克數學家Borůvka做出來了。只是捷克小國寡民，在「布拉格之春」前國際上沒有人關心他們在做些什麼，因此他的研究成果煙沒不彰，直到一九五六年貝耳研究所的Kruskal根據它發表了熟知的「Kruskal 算法」：

「找出最相近而尚未連接的兩點，如果該兩點的連線和網絡中已有的線不構成圈時，則將此線加入網絡中。當網絡中已有$n-1$（n是據點數）線時，構造完成。」



根據Kruskal 算法，首連AB，再連BC，三連DE，四連BD。注意AC雖較DE短但與AB及BC構成圈，故不連。

曾任貝耳研究所通訊理論部門執行總監的Prim提出另一熟知的「Prim 算法」：

先連接最近的兩點，把已經連接的線路看成一個正在茁長的網絡。每次加一條線到這網絡中，所加的線必須是所有可加的線中最短且不構成圈的線。」

按照Prim 算法，上圖五點的連接次序是AB、BC、BD、DE。這兩種算法造出同一棵生成樹並不是偶然的，因爲當最短生成樹是唯一時，任何造最短生成樹的算法必定得到同一結果。

電話公司有了最短生成樹的造法，再以此向私家網絡收費，順理成章，相安無事了十年。

＊＊　　　　＊＊

人心不古、中外皆同。話說北卡州中部有三所頗負盛名的大學，即北卡州大、北卡大和公爵大學；另有一些公司行號依附在這三所大學的勢力範圍內，而統稱爲「研究三角」，在學術界也闖出一些名頭。這三所大學的分佈在一個正三角形的三個頂點上，爲了簡便起見，我們假設此正三角形的邊長是$\sqrt{3}$ 單位。一九六六年研究三角向電話公司申請一張私家網絡把三所大學連接。造這三個據點間的最短生成樹當然用不到什麼Kruskal 、Prim算法，閉著眼睛任取正三角形的兩邊即可，總長爲$2\sqrt{3}$單位。研究三角照章付款，電話公司心滿意足的垂拱收帳。

到了一九六七年，研究三角的私家網絡在正三角的中央添了一個新據點。電話公司更樂不可支，據點愈多，線路愈多，收費愈多。不是嗎？把Kruskal算法拿出一算，最短生成樹是將正三角形的三個頂點連向中央點。根據基本的幾何三角，每一條邊長都是一個單位，所以樹的總長是3個單位。怎麼搞的，樹長從$2\sqrt{3}$ 單位縮短到3個單位了！

此時電話公司疑竇大起，乃派密探到北卡州實地勘察。結果發現該地區並沒有增設第四所大學，正三角形的中央地區是一片沼澤野地，而中心點矗立著一根旗杆，這就是研究三角的第四個據點！看來研究中心果然網羅了高人，隨便豎起一杆就戳了電話公司一個大窟窿。不甘心歸不甘心，規章是規章，不但錢收少了。電話公司還得從三所大學舖設線路連到野地中的旗杆處。

這時消息傳出，在一九六七年這年間美國忽然雨後春筍似的出現了無數的旗杆公司，專門承擔野地立旗杆的工程。電話公司見勢頭不好，趕緊向政府申請另訂收費規章，叫做最短Steiner樹規章。基本上是向行號們討饒，旗杆不用立了，收費就照有旗杆時的最短樹長算（這樣的樹叫做最短Steiner樹）。這樣雖然少收了錢，至少不必往野地拉線路了。

到底少收了多少錢呢？在研究三角的情形下：

$$\frac{3}{2\sqrt{3}}=\frac{\sqrt{3}}{2}\sim .866$$

大約是照八七折收費。會不會更壞呢？貝耳研究所數學中心的Gilbert和中心主任的Pollak作了大量的分析和實驗，得到了很多最短Steiner樹的基本性質，也作了折扣不大於$\sqrt{3}/2$的猜測。

另一方面，解鈴還需繫鈴人。問題從研究三角出，解答也得從那裏找。於是貝耳研究所派了現任副總裁的Mayo 到北卡州去挖人。我在一九六七年夏正在北卡州大寫博士論文。Mayo 說論文擱一擱，你先來再說，我就此進了數學中心做起Steiner問題。

* * * *

一幌二十多年了，我在最短Steiner樹問題上也稍爲作了點貢獻，但在$\sqrt{3}/2$的猜測上却始終一籌莫展。去年年底，電話公司鼓勵工齡二十年以上的員工退休，訂下了優厚的退休章程，我也在合格之列。當時雖心動了一下，但還是爲了最短Steiner樹上的未竟工作而打消了退意，第一是已編好的兩本Steiner樹叢書正在敦促誤期已久的出版商，第二是在爲Algorithmica 雜誌編一期Steiner樹專輯，第三也是不甘放下$\sqrt{3}/2$的猜測。

及後有一頓悟，如果一定要等我證出$\sqrt{3}/2$的猜測才退休，則可能要在貝耳研究所過百歲生日了。解這一難題的關鍵也許不在時間而在「人」！我能作的貢獻是在找一個比我強的人來作此問題。還好，找一個比我強的人倒不是什麼太難的事，我找到了堵丁柱［見附錄］，而堵丁柱在今年四月找到了證明。這個證明並可應用到很多其它的最優化問題上。

* * * *

北卡州野地裏那一面旗在二十餘年後終於徐徐的落下了，我在貝耳研究所的使命也於茲告成。該是重新考慮退休的時候了。感傷嗎？不！因爲仍有許多面舊帆新帆正在胸中乘風欲揚。

參考文獻

有關生成樹部分可參閱 R.L. Graham & P.Hell , On the history of the minimal spanning problem , Ann. History Computing 7 (1985), 43-57。有關Steiner樹部份可參閱F.K.Hwang & D.S.Richards , Steiner tree problems , Networks , 1990 to appear，有關$\sqrt{3}/2$猜測的證明可參閱D.Z.Du & F.K.Hwang, The Steiner ratio conjecture of Gilbert and Pollak is true , Proc. National Acad. Sci., 1990 , to appear 。

附　錄

自由時報在1988年6月21日曾有古威威一篇報導「數學界冒出奇葩，堵丁柱傳奇故事駭聽聞」。這裏就我個人所知略加補充。

堵丁柱祖籍浙江紹興，1948年生於黑龍江齊齊哈爾市。1967年以優異成績自該市重點高中畢業後，正值文革時期全國大專院校都關了門。他遂被分發到一燈泡工廠，一做就十年。1978年三月，學校開了門，他考入東北重機學院大學部一年級。同年夏天，北京科學院十年來第一次招研究生，十年儲存的菁英都一舉來報名投考，光是應用數學所就有二千考生，而只錄取24名，堵丁柱以一個一年級生脫穎而出。1980年我訪問應數所時和他結識。兩個月間我在應數所介紹了一些我多年做不出的問題，對他而言都是前所未聞的，但往往一、二週後他就能提出看法而導至解答。1981年他以師友矚目的成績獲碩士學位，我推薦他入MIT 應數系念博士，系主任Kleitman已首肯，但研究院招生小組却以對大陸院校成績沒有信心爲理由不給他獎學金。堵丁柱乃入加大聖塔芭芭拉分校，三年內以全A成績得博士。時陳省身教授正主持在柏克來甫成立的「數學研究中心」招收博士後生，在七十多名全美各大名校的傑出新博士申請人中挑選六名，堵丁柱又雀屏中選。一年後Kleitman請他去MIT 任助理教授，四年前不願收他爲學生的學校現在却願意把學生交給他！

1987年堵丁柱認爲他在中國可以比在美國做更多事，乃辭去MIT 教職屏擋回國，應數所越過慣例跳過副研究員一級直接聘他爲正研究員。他回國後很受重視，貢獻也多，如擔任全國數學研究補助的評定委員，和程民德、吳文俊、谷超豪、楊樂、馮康等八位數學家共同發起召開「二十一世紀中國數學展望學術討論會」，籌劃主持「組合最優化」的國際會議等；另外也在全國各地演講及寫文章推動學生對數學的興趣。繁忙之餘，他仍做出傑出的研究工作。1989年他是科學院裏獲得青年一等研究獎的唯一數學家。1990年他應邀在Princeton大學訪問一年。

因爲堵丁柱的經歷很特殊，我寫下了這一篇小傳，希望對其他有志數學的青年有所鼓舞。

—本文作者現任職於美國貝耳實驗室—

The End