A. Machì

Dispensa II

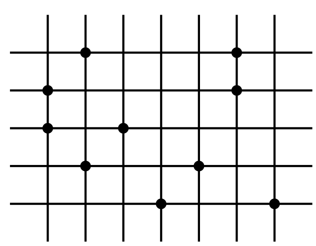
1 Grafi e alberi.

Un *grafo* (non orientato) $G = (V, E)$ consta di un insieme non vuoto V di elementi detti *vertici* o *nodi* e un insieme E di coppie (non ordinate)[1] di vertici dette *archi* o *spigoli* (salvo esplicito avviso parleremo sempre di insiemi V ed E finiti) Se le coppie sono ordinate il grafo si dirà *orientato*. Un arco $e = \{u, v\}$ (o semplicemente $e = uv$) è *incidente* ai vertici u e v, gli *estremi* dell'arco, e si dice *unire* o *congiungere* i due vertici, che allora sono *adiacenti*.[2] Il numero di archi incidenti a un vertice v si dice *grado* del vertice, e si denota con $d(v)$[3].

Teorema 1. *La somma dei gradi di tutti i vertici di un grafo è uguale a due volte il numero degli archi:*

$$\sum_{v \in V} d(v) = 2|E|.$$

Dim. In una tabella come la seguente:



mettiamo in ascissa i vertici e in ordinata gli archi. Segnamo un punto all'incrocio (v, e) se v è incidente a e, e contiamo in due modi diversi, secondo

[1] Se le coppie sono ordinate il grafo si dice *orientato*.
[2] e è l'iniziale dell'inglese "edge".
[3] d è l'iniziale dell'inglese "degree".

cioè le righe verticali e quelle orizzontali, i punti così segnati. Sulla riga sopra un vertice v il numero dei punti segnati dà il numero di archi incidenti a v. Il numero totale dei punti segnati sulle verticali è dunque $\sum_{v \in V} d(v)$. Sulla riga orizzontale corrispondente a un arco e abbiamo due punti segnati (che corrispondono agli estremi dell'arco); contando dunque i punti segnati secondo le righe orizzontali abbiamo due punti per ogni arco, e cosí in totale due volte il numero degli archi. Il risultato segue. $\diamondsuit$

Qualcuno chiama questo teorema "il teorema fondamentale della teoria dei grafi". In particolare, in un qualunque grafo il numero dei vertici di grado dispari è pari.

Se ogni coppia di vertici è congiunta da un arco allora il grafo è il *grafo completo*; si denota con K_n. Se n è il numero dei vertici esso ha $\binom{n}{2}$ archi. Tutti i vertici di K_n hanno grado $n-1$, che è pari se e solo se n è dispari. Dunque, in K_n, o tutti i vertici hanno grado dispari, o tutti hanno grado pari.

Un *cammino* è una successione di vertici $v_1, v_2, \ldots, v_{k-1}, v_k$, dove le coppie $(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)$ sono archi del grafo. Se è possibile andare da un qualunque vertice a ogni altro il grafo è *connesso*; altrimenti esso si spezza in *componenti connesse*. Un cammino $v_1, v_2, \ldots, v_{k-1}, v_k, v_1$, $k > 1$, che comincia e finisce nello stesso vertice e non vi sono ripetizioni tra gli altri vertici è un *ciclo* (o un *circuito*). Se non vi sono cicli il grafo è *aciclico*.

Il teorema che segue mostra come in un grafo connesso il minimo numero di archi è $n-1$ (minimo nel senso che un grafo con meno archi non è più connesso).

Lemma 1. *Un grafo in cui tutti i vertici hanno grado maggiore o uguale a due contiene un ciclo.*

Dim. Partendo da un vertice u qualunque, andiamo lungo un arco uv ad esso incidente; da v prendiamo un arco vw, e cosí via, prendendo sempre un arco mai preso prima. Poiché vi sono solo un numero finito di archi, a un certo punto si arriva a un vertice t tale che tutti gli archi ad esso incidenti sono stati attraversati. Ma essendo t di grado almeno 2, ciò significa che t era già stato visitato, e quindi t fa parte di un ciclo. $\diamondsuit$

Lemma 2. *Un grafo connesso con n vertici ha almeno $n-1$ archi.*

Dim. Per induzione sul numero n dei vertici. Se $n = 1$, il grafo non ha archi e il teorema è ovvio. Sia $n \geq 2$, e sia G un grafo connesso con $n \geq 2$ vertici, e consideriamo il grafo H ottenuto sopprimendo un vertice

v (e quindi anche gli archi ad esso incidenti). H non è necessariamente connesso; siano C_i le sue componenti connesse, $i = 1, 2, \ldots, t$, ciascuna con n_i vertici. Per induzione, C_i ha almeno $n_i - 1$ archi, e dunque in tutto le C_i contengono:

$$(n_1-1)+(n_2-1)+\cdots+(n_t-1) = n_1+n_2+\cdots+n_t-1-1-\cdots-1 = (n-1)-t$$

archi. Ma ciascuna C_i è connessa a v con almeno un arco e dunque reinserendo gli archi soppressi, il grafo G ha almeno $(n-1) - t + t = n - 1$ archi.

$\diamondsuit$

Lemma 3. *Un grafo aciclico con n vertici ha al più $n - 1$ archi.*

Dim. Se $n = 1$, oppure il grafo non ha archi, non c'è niente da dimostrare. Sia G un grafo con $n \geq 2$ vertici. Il grafo ha almeno un vertice v di grado uno, altrimenti avrebbe almeno un ciclo (Lemma 1) contro l'ipotesi. Se sopprimiamo v (e l'arco e ad esso incidente), otteniamo un grafo con $n - 1$ vertici che è ancora aciclico (un ciclo di $G \setminus v$ è anche un ciclo di G). Per induzione questo grafo ha al più $(n-1) - 1 = n - 2$ archi; reinserendo v ed e abbiamo che G ha al più $n - 1$ archi. $\diamondsuit$

Teorema 2. *Sia G un grafo con n vertici. Le seguenti proposizioni sono equivalenti:*

i) G è connesso e aciclico;
ii) G è aciclico e ha $n - 1$ archi;
iii) G è connesso e ha $n - 1$ archi;
iv) G è connesso e la soppressione di un arco disconnette il grafo;
v) due vertici qualunque sono congiunti da uno e un solo cammino;
vi) G è aciclico, ma l'aggiunta di un arco crea un ciclo e uno solo.

Dim. Per $n = 1$ non c'è niente da dimostrare. Sia $n \geq 2$.

i) $\Rightarrow$ *ii)* Lemmi 1 e 2.

ii) $\Rightarrow$ *iii)* Se G non è connesso, sia $G = G_1 \cup G_2 \cup \ldots \cup G_k$ la decomposizione in componenti connesse. Poiché G è aciclico, lo stesso accade per i G_i. Se $|G_i| = n_i$, per *vi)* G_i ha $n_i - 1$ archi e quindi G ne ha $\sum_{i=1}^{k}(n_i - 1) = \sum_{i=1}^{k} n_i - k = n - k$. Se $k > 1$, cioè se G non è connesso, G ha meno di $n - 1$ archi, contro l'ipotesi.

iii) $\Rightarrow$ *iv)* Sopprimendo un arco si ottiene un grafo con $n - 2$ archi, che per il Teor. 2 non è connesso.

iv) $\Rightarrow$ *v)* Essendo il grafo connesso, due vertici qualunque sono congiunti da un cammino. Se vi sono due cammini da u a v, $u = v_1, v_2, \ldots, v_{k-1}, v_k = v$

e $u = u_1, u_2, \ldots, u_{k-1}, v_k = v$ si ha il ciclo

$$u = v_1, v_2, \ldots, v_{k-1}, v_k, u_{k-1}, u_2, u_1 = u,$$

e sopprimendo un arco di questo ciclo il grafo resta connesso.

v) $\Rightarrow$ *vi)* Se vi fosse un ciclo, due vertici di questo ciclo sarebbero congiunti da almeno due cammini, contro l'ipotesi. Aggiungendo un arco $e = (u, v)$, poiché c'è già un cammino tra u e v si crea un ciclo. E se ne crea uno solo. Se infatti si creassero due cicli, poiché allora questi contengono entrambi e, sopprimendo e si vede che u e v sono congiunti in G dai due cammini che costituiscono le parti dei due cicli che non contengono e.

vi) $\Rightarrow$ *i)* Se G non è connesso, aggiungendo un arco che unisce due vertici in due componenti diverse non si creano cicli. Se si creano due cicli, l'arco aggiunto fa parte di entrambi, e la sua soppressione mostra ceh il grafo di partenza conteneva un ciclo. $\diamondsuit$

Un grafo che soddisfa una delle condizioni equivalenti del Teor. 2 si chiama *albero*, e si denota con T (iniziale dell'inglese *tree*).

Corollario. *In un albero con più di un vertice esistono almeno due vertici di grado* 1 (*vertici estremi o foglie*).

Dim. Consideriamo un cammino p di lunghezza massima, e siano u e v i suoi vertici estremi. Se c'è un arco e incidente a u, e ad alcun altro vertice di p, il cammino $p \cup e$ è più lungo di p, contro la massimalità di p. Allora e è incidente ad un vertice w che precede u in p, e in tal caso il cammino $w, \ldots, u, w$ è un ciclo. $\diamondsuit$

Sia G un grafo connesso. Se G non ha cicli, allora è un albero (Teor. 3, *i*). Altrimenti, scegliamo un ciclo e sopprimiamo un arco di questo; il grafo resta connesso. Continuando il procedimento si resta con un albero che connette tra loro tutti i vertici di G. Di un tale albero si dice che *ricopre* il grafo ("spanning tree" in inglese).

Una procedura per trovare uno spanning tree in un grafo connesso è la "visita in profondità" DFS (Depth-First Search), che procede come segue:

scegliere un vertice a caso e costruire un cammino a partire da questo finché non si può più proseguire senza ripetere un vertice già visitato (searched); il vertice dove questo cammino si ferma sarà una foglia dell'albero. Si risalga allora (*backtracking*) *fino ad un vertice a partire dal quale si può costruire un nuovo cammino.*

La caratteristica di questo metodo è che il nuovo vertice viene scelto tra quelli adiacenti all'ultimo dei vecchi vertici. Il metodo si può anche usare per vedere se un dato grafo è connesso. Se infatti tutti i vertici sono raggiunti nella DFS il grafo è connesso, altrimenti non lo è.

La DFS si può descrivere come segue: si costruisce una lista di vertici, e a ogni passo la si svuota a cominciare dal vertice che al momento attuale si trova in coda alla lista (*last–in, first–out*). Una tale lista viene anche detta *pila* (*stack*, in inglese). Un vertice viene aggiunto alla pila quando si avanza nella visita, e si rimuove quando si backtrack da esso.

PROCEDURA **DFS**

INIZIO *pila* P=$\{x\}$*;*
finché P *non è vuota fare*
  u:= cima della pila;
    per ogni vertice v adiacente a u fare:
      se u non è stato ancora visitato,
      aggiungere u in cima alla pila;
    altrimenti fare:
  rimuovere u dalla pila
FINE

Se i vertici del grafo sono etichettati $1, 2, \dots, n$, si può effettuare una DFS cominciando con 1, e ad ogni passo prendere l'arco che congiunge il vertice attuale con il vertice ad esso adiacente e che ha l'indice più basso. Se il vertice attuale non è adiacente a un nuovo vertice, si torna indietro, fino a un vertice a partire dal quale si può iniziare un nuovo cammino. L'albero ottenuto con la procedura DFS ha un aspetto di albero "lungo e rado".

Un altro metodo per costruire un albero è la "visita in larghezza" BFS (Breadth-First Search). Contrariamente alla DFS, nella quale si passa da un vertice a uno nuovo finché è possibile, nella BFS si visitano tutti i vertici adiacenti a quello attuale prima di passare a uno nuovo. Non c'è quindi backtracking. La lista di vertici che si costruisce è una *coda* (*queue*, in inglese): essa viene svuotata a ogni passo a cominciare dal vertice che al momento attuale si trova in cima alla lista (*first–in, first–out*).

PROCEDURA **BFS**

INIZIO *coda* $Q = \{v\}$
finché Q non è vuota fare:
  w: fronte della coda;
    se w è adiacente a un nuovo vertice u fare:
      aggiungere u alla fine della coda;
    altrimenti fare:

rimuovere w dalla coda.
FINE

L'albero BFS è "basso e folto".

2 Alberi di copertura minimali

Se a ogni arco e del grafo è associato un numero reale non negativo $w(e)$ (w sta per "weight", *peso* dell'arco; invece del peso si può pensare anche a un *costo*) nel cercare un albero che ricopre il grafo ha senso cercarne uno di peso minimo, cioè un albero nel quale è minima la somma dei pesi di tutti gli archi. Diremo per brevità "albero minimo" invece di "albero di peso minimo", e lo indicheremo con la sigla MST (M*inimum* S*panning* T*ree*).

L'idea generale degli algoritmi di ricerca di un MST è di disporre a ogni passo di un insieme A di archi che è un sottoinsieme di un MST, e di aggiungere ad A un arco e in modo che $A \cup \{e\}$ sia ancora un sottoinsieme di un MST.

In questo paragrafo presentiamo due algoritmi per determinare un MST, uno dovuto a J. B. Kruskal (1956), l'altro a R.C. Prim (1957). Sono entrambi algoritmi "greedy" (a ogni passo si cerca la migliore soluzione immediata).

Nell'algoritmo di Kruskal l'insieme A è una *foresta* (una unione di alberi disgiunti), e l'arco e che si aggiunge ad A è un arco del grafo di peso minimo (tale che A sia ancora una foresta, cioè non si formino cicli). All'inizio A è un arco di peso minimo tra tutti gli archi del grafo.

Algoritmo di Kruskal

- INIZIO A:= arco di peso minimo.
- RIPETERE la seguente operazione: aggiungere ad A un arco di peso minimo, evitando la creazione di cicli.
- FINE

Nell'algoritmo di Prim l'insieme A è un albero, e l'arco e da aggiungere ad A è un arco minimo che unisce un vertice di A con uno non contenuto in A. Quindi a ogni passo A è un albero; all'inizio A consta di un solo vertice.

Algoritmo di Prim

- INIZIO A:= $\{v\}$ (un vertice qualunque).

- RIPETERE la seguente operazione: aggiungere ad A un arco minimo tra quelli incidenti a un vertice di A evitando la creazione di cicli
- FINE

Che i due algoritmi diano effettivamente alberi di peso minimo è una conseguenza del seguente Lemma.

Lemma. *Se $(S,\ V \setminus S)$ è una partizione dei vertici del grafo, ogni arco di peso minimo tra quelli che uniscono un vertice di S con uno di $V \setminus S$ è contenuto in un MST.*

Dim. Sia T un MST, e tra gli archi con i vertici uno in S e l'altro in $V \setminus S$ sia $e = (u, v) \notin T$ uno di peso minimo. In T c'è un cammino che unisce u a v, e dunque un arco e' di T tra un vertice di S e uno di $V \setminus S$. Se aggiungiamo e a T otteniamo un grafo che ha un ciclo, e sopprimendo e' un albero. Avendosi $w(e) \leq w(e')$, se $w(e) < w(e')$ sopprimendo e' si ottiene un albero di peso minore di T, escluso, per cui $e \in T$. Se $w(e) = w(e')$, allora sopprimendo e' abbiamo l'albero $(T \setminus \{e'\}) \cup \{e\}$ che ha lo stesso peso di T (e quindi è minimo) e contiene e. ♢

In particolare, se tra due insiemi di vertici S e $V \setminus S$ c'è un arco di peso inferiore a quello di tutti gli archi che uniscono i due insiemi, questo arco appartiene a ogni MST.

Corollario. *Sia A un sottoinsieme di archi che è parte di un MST, $A \subseteq T$, e sia $(S, V \setminus S)$ una partizione dei vertici, dove S contiene tutti i vertici di A. Se $e = (u, v) \notin T$ è un arco tale che $u \in S$ e $v \in V \setminus S$ ed è di peso minimo, allora anche $A \cup \{e\}$ è parte di un MST.*

Dim. L'aggiunta di e non può creare un ciclo perché altrimenti vi sarebbe un cammino in A che comincia in u e finisce in v e quindi un arco (x, y) di questo cammino di A con $x \in S$ e $y \in V \setminus S$, mentre per ipotesi i vertici di A sono tutti contenuti in S. Con la partizione $(S, V \setminus S)$ il risultato segue ora dal lemma. ♢

Nel caso dell'algoritmo di Kruskal, una volta costruito A come parte di un MST vi sono due casi:

i) aggiungiamo un arco e che unisce due vertici isolati u e v e prendiamo per S l'insieme dei vertici di A più u. Per il corollario anche $A \cup \{e\}$ è parte di un MST.

ii) aggiungiamo un arco con un vertice in A, e prendiamo per S l'insieme dei vertici di A. Per il corollario anche $A \cup \{e\}$ è parte di un MST.

Nell'algoritmo di Prim, a ogni passo si aggiunge un arco di peso minimo tra quelli incidenti ai vertici di A: siamo quindi nel caso *ii*) ora visto. Si osservi che ad ogni passo A è un albero, parte di un MST.

I due algoritmi non hanno la stesa complessità computazionale. Si può dimostrare che, se G è un grafo con n vertici e m archi, l'algoritmo di Kruskal costruisce un albero di copertura di peso minimo in $O(m \log m)$ passi, mentre quello di Prim ne richiede $O(n^2)$. Quindi, se G è un grafo *denso* (cioè il numero di archi è dell'ordine di n^2), è più efficiente l'algoritmo di Prim, che impiega $O(n^2)$ passi anziché $O(n^2 \log n)$; se G è *sparso* (cioè $m \sim n$), allora è più efficiente quello di Kruskal.

3 Grafi euleriani e hamiltoniani

Un grafo si dice *euleriano* se esiste un circuito che passa per tutti gli archi e una sola volta.

In quali grafi esiste un tale cammino? Intanto il grafo deve essere connesso; inoltre, un circuito entra ed esce da ciascun vertice uno stesso numero di volte, e quindi i vertici devono avere tutti grado pari. Vi sono quindi due condizioni necessarie: *i*) il grafo deve essere connesso, *ii*) i vertici devono avere tutti grado pari. Queste due condizioni sono anche sufficienti, come dimostra il teorema che segue.

Teorema 3. *Condizione necessaria e sufficiente affinché un grafo ammetta un circuito euleriano è che sia connesso e tutti i vertici abbiano grado pari.*

Dim. La necessità è già stata dimostrata. Per la sufficienza, supponiamo di cominciare un cammino in un vertice u, e di proseguire finché sia possibile uscendo da un vertice seguendo sempre un arco non percorso prima. Questo procedimento non può proseguire indefinitamente (vi è soltanto un numero finito di vertici e archi), ma dato che in ogni vertice sono incidenti un numero pari di archi, una volta entrati in un vertice vi è sempre un arco da cui uscire, tranne che per il vertice iniziale u. Allora il cammino termina in u, e si ottiene cosí un circuito $\mathcal{C}$. Se tutti gli archi sono stati percorsi abbiamo trovato un circuito euleriano. Altrimenti, esistono archi non appartenenti a $\mathcal{C}$ e quindi almeno un vertice v di $\mathcal{C}$ incidente ad archi non appartenenti a $\mathcal{C}$ (altrimenti $\mathcal{C}$ sarebbe una componente connessa diversa da tutto il grafo, contro il fatto che il grafo è connesso). Poiché vi è un numero pari di archi nel circuito $\mathcal{C}$ incidenti a v, e v è di grado pari, esistono un numero pari di archi non appartenenti a $\mathcal{C}$ (cioè non percorsi prima) e incidenti a v, e lo stesso accade per tutti i vertici incidenti ad archi non percorsi. Cominciamo allora un nuovo cammino partendo da v e seguendo solo archi non percorsi prima, e analogamente a quanto visto sopra otteniamo un circuito $\mathcal{C}'$ che termina in v. A partire da u possiamo allora costruire un circuito più lungo unendo $\mathcal{C}$ e $\mathcal{C}'$ in v. Se l'intero grafo non è stato esaurito, si può ampliare ulteriormente il circuito nello stesso modo fino a ottenere un circuito euleriano. $\diamondsuit$

Se in un grafo esiste un cammino da un vertice u a un vertice $v \neq u$ percorrendo una sola volta tutti gli archi, allora il cammino segue un certo arco incidente a u per eventualmente rientrarvi e ripartirne un certo numero di volte. Poiché il cammino non termina in u, $d(u)$ è dispari; analogamente $d(v)$ è dispari, e gli altri vertici sono di grado pari. Abbiamo allora:

Teorema 4. *Condizione necessaria e sufficiente affinché un grafo ammetta un cammino da un vertice u a un vertice v percorrendo una sola volta tutti gli archi è che u e v siano i soli vertici di grado dispari.*

Dim. La necessità è già stata dimostrata. Per la sufficienza, aggiungiamo un arco uv; tutti i vertici sono ora di grado pari, e per il Teor. 1 esiste un circuito euleriano. Sopprimendo l'arco uv si resta con un cammino da u a v come richiesto. $\diamondsuit$

La nozione analoga a quello di circuito euleriano, ma che riguarda i vertici è, quella di circuito *hamiltoniano.* Un circuito hamiltoniano è un circuito che tocca tutti i vertici una sola volta. Un cammino hamiltoniano è un cammino che tocca tutti i vertici una sola volta.

Un *grafo hamiltoniano* è un grafo che ammette un circuito hamiltoniano. Contrariamente al caso euleriano (nel quale basta controllare che ogni vertice abbia grado pari), stabilire se un grafo è hamiltoniano è molto più difficile (si tratta di un problema NP–completo). Esistono vari risultati che danno condizioni sufficienti per l'esistenza di un circuito hamiltoniano in un grafo dato; uno dei più noti è il teorema seguente, dovuto a G.A. Dirac.

Teorema 5. *Se G è un grafo su n vertici in cui ogni vertice ha grado almeno $n/2$, allora G è hamiltoniano.*

Dim. Supponiamo per assurdo che per qualche n esista un grafo su n vertici che soddisfa la condizione dell'enunciato ma non ammette un circuito hamiltoniano. Tra i grafi su n vertici con queste proprietà sia G uno con il massimo numero di archi; quindi, se aggiungiamo un arco a G, il nuovo grafo sarà hamiltoniano.

Siano u e v due vertici non adiacenti di G (perché si possono sempre trovare due vertici non adiacenti?). Nel grafo ottenuto da G aggiungendo l'arco $\{u,v\}$, per quanto detto, c'è un circuito hamiltoniano; rimuovendo questo arco, otteniamo in G un cammino hamiltoniano da u a v. Sia questo cammino $u = x_1,\, x_2,\, \ldots,\, x_{n-1},\, x_n = v$.

Definiamo i due insiemi $U = \{i : u \text{ è adiacente a } x_{i+1}\}$ e $V = \{i : v \text{ è adiacente a } x_i\}$ (V è l'insieme degli indici dei vicini di v, mentre U è l'insieme degli indici dei "predecessori" dei vicini di u). Entrambi questi insiemi sono sottoinsiemi di $\{1, 2, \ldots, n-1\}$ e, per l'ipotesi sui gradi, hanno almeno $n/2$ elementi. Quindi U e V devono avere almeno un elemento in comune; sia i_0 un tale elemento.

Possiamo allora costruire il circuito $u = x_1,\, x_2,\, \ldots,\, x_{i_0},\, v = x_n,\, x_{n-1},\, x_{n-2},\, \ldots,\, x_{i_0+1},\, u$. Questo circuito passa per ogni vertice una sola volta, contraddicendo l'ipotesi che G non fosse hamiltoniano. $\diamondsuit$

4 Alberi etichettati

Due alberi si dicono *isomorfi* se esiste una corrispondenza biunivoca f tra i vertici che conserva gli archi: se (u,v) è un arco, anche $(f(u), f(v))$ lo è. Si vede facilmente che, a meno di isomorfismi, su 1 e 2 vertici c'e un solo albero, e che su 3, 4 e 5 vertici vi sono rispettivamente, 1, 2 e 3 alberi:

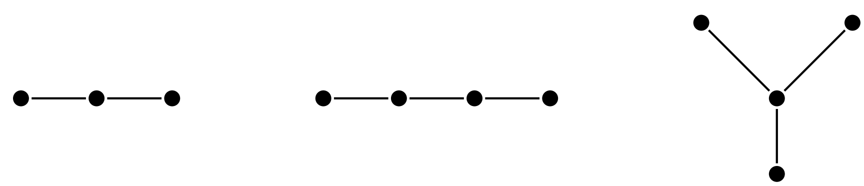


Fig. 1 I tre alberi su 3 e 4 vertici a meno di isomorfismi.

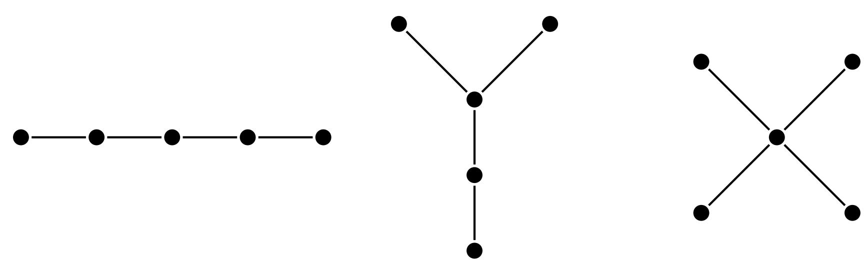


Fig. 2 I tre alberi su 5 vertici a meno di isomorfismi

La determinazione del numero di alberi su n vertici a meno di isomorfismi è un problema estremamente difficile; si conoscono formule asintotiche ma non formule chiuse. Non è difficile invece determinare il numero di alberi *etichettati* su n vertici. Un albero etichettato è un albero nel quale ai vertici sono associate etichette (di solito le cifre 1,2,...). Due alberi T e T' su n vertici sono diversi se per almeno una coppia di vertici etichettati (i, j) l'arco che unisce i e j esiste in T ma non in T'. Gli alberi etichettati su 3, 4 e 5 vertici sono allora sono 3,16 e 125, rispettivamente, come dimostra il Teor. 5 qui sotto.

Se consideriamo il grafo completo K_n su n vertici etichettati 1,2,...n, contare gli alberi etichettati su n vertici equivale a contare gli alberi che ricoprono questo K_n.

Teorema 6. (Cayley, 1889) *Il numero di alberi etichettati su n vertici è* n^{n-2}.

Daremo due dimostrazioni di questo teorema. Altre due, che fanno uso di una formula di Abel, si trovano nell'Appendice a questa Dispensa n. 2.

i) Prima dimostrazione (Prüfer, 1918). Stabiliamo una corrispondenza biunivoca tra gli alberi etichettati su n vertici $1, 2, \dots, n$ e le $(n-2)-$ple $(i_1, i_2, \dots, i_{n-2})$, dove gli i_k sono interi da 1 a n, eventualmente con ripetizioni. Poiché queste ultime sono in numero di n^{n-2} si avrà il risultato. Chiameremo la $(n-2)-$pla associata all'albero T *codice di Prüfer* di T.

Scriviamo la dimostrazione sotto forma di algoritmo:

Albero etichettato $\longrightarrow$ Codice di Prüfer

- INIZIO La lista vuota **P**.
- RIPETERE la seguente operazione:

  se v è il vertice estremo con l'etichetta minima, aggiungere a **P** l'etichetta dell'unico vertice unito a v ed eliminare dall'albero sia v che l'arco al quale v appartiene.

  FINCHE' non rimane un solo arco.
- La lista **P** è il codice di Prüfer dell'albero.

Si noti che il numero di volte in cui un vertice v compare nel codice è $d(v)-1$.

Codice di Prüfer $\longrightarrow$ Albero etichettato

- Siano **P** il codice di Prüfer ed **L**$= \{1, 2, \dots, n\}$ la lista delle etichette.

- FINCHE' la lista **P** non è vuota, ripetere la seguente operazione:

  sia p il primo elemento della lista **P**, e sia s il più piccolo elemento di **L** che non compare in **P**. Unire p a s con un arco ed eliminare p da **P** e s da S

- Quando P è vuota, restano due elementi in L; congiungendo i vertici corrispondenti si completa l'albero.

È facile vedere che se si parte da un albero etichettato, si determina la $(n-2)-$pla corrispondente, e poi l'albero corrispondente a questa $(n-2)-$pla, si ottiene l'albero di partenza, e quindi la corrispondenza biunivoca richiesta.

Diamo ora la seconda dimostrazione del teorema di Cayley.

Seconda dimostrazione. In questa dimostrazione contiamo gli alberi etichettati su n vertici nei quali un vertice fissato v ha grado k. Sia $T(n,k)$ questo numero; sommando su k da 1 a $n-1$ si otterrà il risultato. Sia A un albero nel quale v ha grado k, e sia $e=(w,u)$ un arco non incidente a v. Sopprimendo e otteniamo due alberi, uno contenente v e uno tra w e u, diciamo w, e l'altro u. Se ora uniamo v e u otteniamo un albero B nel quale $d(v)=k+1$. Contiamo ora in due modi diversi le coppie (A,B) ottenute come sopra.

i) A si può scegliere in $T(n,k)$ modi, e poiché B è univocamente determinato dall'arco e, che si può scegliere in $(n-1)-k$ modi, per ogni scelta di A abbiamo $n-1-k$ alberi B, e dunque le coppie (A,B) sono in tutto $T(n,k)(n-1-k)$.

ii) Sia B un albero nel quale $d(v)=k+1$, e siano $B_1, B_2, \ldots, B_{k+1}$ gli alberi ottenuti da B sopprimendo il vertice v e gli archi $e_i=(v,u_i)$ ad esso incidenti. Sopprimendo da B uno degli archi e_i e unendo u_i a un qualunque vertice non appartenente a B_i si ottiene un albero A nel quale v ha grado k.

L'albero B si può scegliere in $T(n,k+1)$ modi. Per ciascuna scelta di uno di questi, il vertice u_i unito a v si può scegliere in $k+1$ modi, e per ciascuna scelta di u_i il vertice diverso da v non appartenente a B_i si può scegliere in $(n-1)-n_i$ modi. Ne segue che l'albero A si può scegliere in $\sum_{i=1}^{k+1}(n-1-n_i)$ modi, ovvero essendo $\sum_{i=1}^{k+1} n_i = n-1$, in $(n-1)k$ modi.

Abbiamo quindi:

$$(n-k-1)T(n,k) = (n-1)kT(n,k+1)$$

da cui

$$T(n,k) = \frac{(n-1)k}{(n-k-1)}T(n,k+1).$$

Analogamente:

$$T(n,k+1)=\frac{(n-1)(k+1)}{(n-k-2)}T(n,k+2),$$

da cui

$$T(n,k)=\frac{(n-1)k}{(n-k-1)}\cdot\frac{(n-1)(k+1)}{(n-k-2)}\cdot T(n,k+2)$$

e iterando:

$$T(n,k)=\frac{(n-1)k}{(n-k-1)}\cdot\frac{(n-1)(k+1)}{(n-k-2)}\cdots\frac{(n-1)(n-2)}{1}\cdot T(n,n-1).$$

Mettendo in evidenza $n-1$ e ricordando che $T(n,n-1)=1$:

$$\begin{aligned} T(n,k) &= \frac{(n-2)\cdots(k+1)\cdot k}{(n-k-1)(n-k-2)\cdots 1}(n-1)^{n-k-1} \\ &= \binom{n-2}{n-k-1}\cdot(n-1)^{n-k-1} \\ &= \binom{n-2}{k-1}\cdot(n-1)^{n-k-1}. \end{aligned}$$

Sommando per k che va da 1 a $n-1$ otteniamo:

$$T(n)=\sum_{k=1}^{n-1}T(n,k)=\sum_{k=1}^{n-1}\binom{n-2}{k-1}\cdot(n-1)^{n-k-1}$$

e posto $k-1=h$:

$$T(n)=\sum_{h=0}^{n-2}\binom{n-2}{h}\cdot(n-1)^{n-2-h}\cdot 1^h=[(n-1)+1]^{n-2}=n^{n-2}.$$

come si voleva. $\diamondsuit$

Esempio. L'albero su $V=\{1,2,\ldots,10\}$ i cui archi sono:

$$\{1,6\},\{1,8\},\{1,9\},\{2,8\},\{3,7\},\{3,10\},\{4,9\},\{5,10\},\{7,8\}$$

ha codice di Prüfer (8,9,10,1,1,8,7,3).

Esercizi

1. In un albero su n vertici la somma dei gradi è $2n-2$.

2. Quanti alberi etichettati vi sono su 5 vertici 1,2,3,4,5 tali che $d(1)=3, d(2)=2, d(3)=d(4)=d(5)=1$?

3. In un grafo con 8 archi nel quale tutti i vertici hanno grado 4, quanti vertici vi sono?

4. Se in una riunione di persone queste si stringono l'un l'altra la mano, vi sono almeno due persone che stringono lo stesso numero di mani.

5. Il numero di foglie di un albero è maggiore o uguale del massimo tra i gradi dei vertici.

6. Sia $n \geq 2$. Determinare quanti degli n^{n-2} alberi etichettati su n vertici hanno:

i) un vertice di grado $n-1$;

ii) un vertice di grado $n-2$.

7. Una foresta con $n-1$ archi è un albero (una *foresta* è un'unione disgiunta di alberi).

Appendice

In questa Appendice diamo una dimostrazione della formula di Abel che generalizza la formula del binomio di Newton. Dando valori particolari ai parametri che vi compaiono si otterrà una formula che, assieme a un modo di contare gli alberi etichettati su n vertici diverso da quelli visti finora, permetterà una nuova dimostrazione del teorema di Cayley.

Dimostriamo dapprima, con la tecnica di Abel, la formula del binomio di Newton:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k x^{n-k} \tag{1}$$

Scriviamo la (1) come un polinomio in x:

$$(x+y)^n = x^n + \binom{n}{1} y x^{n-1} + \binom{n}{2} y^2 x^{n-2} + \cdots + \binom{n}{n} y^n \tag{2}$$

Per $n = 1$ la (2) è chiara. Sia allora $n > 1$; supponiamo la (2) vera per n e dimostriamola vera per $n+1$. Moltiplichiamo entrambi i membri della (2) per $n+1$ e integriamo rispetto a x; otteniamo:

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n} \frac{n+1}{n+1-k} \binom{n}{k} y^k x^{n+1-k} + c.$$

Per $x = 0$ si trova $c = y^{n+1}$, e dunque questo è il valore della costante c. Ne segue:

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n} \frac{(n+1)}{n+1-k} \binom{n}{k} y^k x^{n+1-k} + y^{n+1} \tag{3}$$

Ma:

$$\frac{(n+1)}{n+1-k} \binom{n}{k} = \frac{(n+1)}{n+1-k} \frac{n!}{k!(n-k)!} = \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k}.$$

La (3) diventa allora:

$$\begin{aligned} (x+y)^{n+1} &= x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} y^k x^{n+1-k} + y^{n+1} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} y^k x^{n+1-k}, \end{aligned}$$

come si voleva.

La formula di Abel è ora la seguente:

$$(x+y)^n = x^n + y\sum_{k=1}^{n}\binom{n}{k}(y-kz)^{k-1}(x+kz)^{n-k} \tag{4}$$

(si ottiene la (1) per $z=0$). Dimostriamola, come la (1) per induzione su n. Per $n=1$ il risultato è chiaro. Sia $n>1$ Come sopra, moltiplichiamo per $n+1$ e integriamo rispetto a x; si ottiene:

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}y(y-kz)^{k-1}(x+kz)^{n-k} + c \tag{5}$$

(il simbolo binomiale $\binom{n+1}{k}$ è ottenuto come prima). Facciamo vedere che il valore della costante c è il termine per $k=n+1$ della (4) cioè $y(y-(n+1)z)^n$. Poniamo $x=-(n+1)z$ nella (4):

$$(y-(n+1)z)^n = (-(n+1)z))^n + \sum_{k=1}^{n}\binom{n}{k}y(y-kz)^{k-1}([k-(n+1)]z)^{n-k}, \tag{6}$$

e nella (5):

$$\begin{aligned}(y-(n+1)z)^{n+1} &= \\ (-(n+1)z))^{n+1} &+ \sum_{k=1}^{n}\binom{n+1}{k}y(y-kz)^{k-1}([k-(n+1)]z)^{n+1-k} + c.\end{aligned} \tag{7}$$

Dimostriamo che i termini a secondo membro della (6) moltiplicata per $(n+1)z$ sono uguali a quelli della (7) (escluso c) con il segno cambiato. Si ha intanto:

$$(-(n+1)z))^n(n+1)z = (-1)^n(n+1)^n z^n(n+1)z = (-1)^n(n+1)^{n+1}z^{n+1}.$$

Inoltre il termine generico della (5) moltiplicata per $(n+1)z$:

$$\frac{n!}{k!(n-k)!}(y-kz)^{k-1}([k-(n+1)]z)^{n-k}(n+1)z$$

è uguale a:

$$\frac{(n+1)!}{k!(n-k)!}(y-kz)^{k-1}(-1)^{n-k}(n+1-k)^{n-k}z^{n+1-k}. \tag{8}$$

Il termine generico della (6) vale:

$$\frac{(n+1)!}{k!(n+1-k)!}(y-kz)^{k-1}(-1)^{n+1-k}(n+1-k)^{n+1-k}z^{n+1-k},$$

ovvero, essendo $(n+1-k)! = (n+1-k)(n-k)!$,

$$\frac{(n+1)!}{k!(n-k)!}(y-kz)^{k-1}(-1)^{n+1-k}\frac{(n+1-k)^{n+1-k}}{n+1-k}z^{n+1-k},$$

cioè

$$\frac{(n+1)!}{k!(n-k)!}(y-kz)^{k-1}(-1)^{n+1-k}(n+1-k)^{n-k}z^{n+1-k},$$

che è l'opposto della (8).

Sommando allora la (6) moltiplicata per $(n+1)z$ e la (7) questi termini si annullano e si trova:

$$c = (y-(n+1)z)^n(n+1)z + (y-(n+1)z)^{n+1}.$$

Mettendo in evidenza $(y-(n+1)z)^n$:

$$c = (y-(n+1)z)^n((n+1)z + y - (n+1)z) = y(y-(n+1)z)^n.$$

Questo valore di c è il termine per $k = n+1$ della

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n+1}\binom{n+1}{k}y(y-kz)^{k-1}(x+kz)^{n+1-k}$$

e ciò conclude la dimostrazione.

Per dimostrare il teorema di Cayley trasformiamo la (4) come segue. Deriviamola rispetto a y:

$$n(x+y)^{n-1} = \sum_{k=1}^{n}\binom{n}{k}(y-kz)^{k-1}(x+kz)^{n-k} + y(\sum\cdots)'.$$

Posto $x = n, y = 0, z = -1$ abbiamo:

$$n\cdot n^{n-1} = \sum_{k=1}^{n}\binom{n}{k}k^{k-1}(n-k)^{n-k} = \sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k} + n^{n-1},$$

da cui:

$$(n-1)n^{n-1} = \sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k}. \tag{9}$$

Trasformiamo ora questa uguaglianza. Il secondo membro si scrive:

$$\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k} = \sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1}(n-k) =$$

$$= n\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1} - \sum_{k=1}^{n-1}\binom{n}{k}k^{k}(n-k)^{n-k-1},$$

e avendosi $\binom{n}{k} = \binom{n}{n-k}$, possiamo sostituire nell'ultima somma k con $n-k$:

$$= n\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1} - \sum_{k=1}^{n-1}\binom{n}{k}(n-k)^{n-k}k^{k-1}.$$

Portiamo la seconda somma a primo membro; otteniamo:

$$2\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k} = n\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1}.$$

Dalla (9) segue allora:

$$2(n-1)n^{n-1} = n\sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1},$$

da cui:

$$2(n-1)n^{n-2} = \sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1}. \tag{10}$$

Per applicare la (10) contiamo gli alberi come segue. Dato un albero su n vertici, togliendo un arco esso si spezza in due alberi, su k e $n-k$ vertici, $k = 1, 2, \ldots, n-1$. Viceversa, per ogni partizione degli n vertici in k e $n-k$, ciò che si può fare in $\binom{n}{k}$ modi, costruendo tutti gli alberi su ciascuno dei due sottoinsiemi e congiungendo uno dei k vertici con uno degli $n-k$ in tutti i $k(n-k)$ modi possibili si ottengono $\binom{n}{k}k(n-k)T(k)T(n-k)$ alberi. Sommando da 1 a $n-1$ ogni partizione in k e $n-k$ vertici si ottiene due volte, una volta quando si scelgono k vertici, la seconda quando se ne scelgono $n-k$. Infine, potendo togliere $n-1$ archi, ogni albero viene ripetuto $n-1$ volte. Ne segue:

$$2(n-1)T(n) = \sum_{k=1}^{n-1}\binom{n}{k}k(n-k)T(k)T(n-k). \tag{11}$$

Per induzione, $T(k) = k^{k-2}$ e $T(n-k) = (n-k)^{n-k-2}$; con questi valori la (11) diventa:

$$2(n-1)T(n) = \sum_{k=1}^{n-1}\binom{n}{k}k^{k-1}(n-k)^{n-k-1}, \tag{12}$$

che confrontata con la (10) dà:

$$2(n-1)T(n) = 2(n-1)n^{n-2},$$

da cui $T(n) = n^{n-2}$.

Vediamo ora come ritrovare lo stesso risultato attribuendo particolari valori ai parametri che compaiono nella formula di Abel (4). Sostituendo in questa formula

$$n, \quad k, \quad y, \quad z, \quad x$$

rispettivamente con:

$$n-2, \;\; k-1, \;\; 1, \;\; -1, \;\; n-1$$

essa diventa:

$$n^{n-2} = \sum_{k=1}^{n-1} \binom{n-2}{k-1} k^{k-2}(n-k)^{n-k-1}. \tag{13}$$

Per induzione, $k^{k-2} = T(k)$ e $(n-k)^{n-k-1} = (n-k)^{n-k-2} \cdot (n-k) = T(n-k) \cdot (n-k)$, e dunque il secondo membro della (13) si può scrivere:

$$\sum_{k=1}^{n-1} \binom{n-2}{k-1} (n-k)T(k)T(n-k). \tag{14}$$

Facciamo vedere che questo numero è precisamente $T(n)$. Fissiamo due vertici u e v, un sottoinsieme di k vertici che contiene u ma non v e due alberi, uno sui k vertici scelti e uno sugli $n-k$ rimanenti. Unendo u con uno di questi ultimi si ottiene un albero sugli n vertici dati. Con questa operazione, per ogni scelta di $k-1$ vertici $u_1, u_2, \ldots, u_{k-1}$ distinti da u e v, a partire dai $T(k)$ alberi sui k vertici $u_1, u_2, \ldots, u_{k-1}, u$, dai $T(n-k)$ sui rimanenti e dagli $n-k$ archi che uniscono u e uno degli $n-k$ vertici, si ottengono $T(k)T(n-k)(n-k)$ alberi su n vertici. Tra gli $n-2$ vertici diversi da u e v un sottoinsieme di k vertici contenente u si può scegliere in $\binom{n-2}{k-1}$ modi. Per ognuna di queste scelte abbiamo quindi $T(k)T(n-k)(n-k)$ alberi su n vertici, e sommando queste scelte per k che va da 1 a $n-1$ si ottiene la (14). Ogni albero si ottiene in questo modo: fissato infatti un albero sugli n vertici dati, si consideri il cammino da v a u. Sopprimendo l'arco incidente a u si ottengono due alberi su k e $n-k$ vertici, per un certo k; dunque uno degli alberi visti sopra.