Azzolini Riccardo 2019-03-13

Alberi

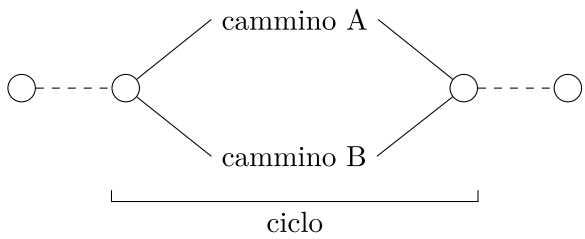
1 Albero

Un **albero** ha tre definizioni equivalenti:

1. un grafo non orientato connesso e aciclico;
2. un grafo non orientato che ammette uno e un solo cammino tra qualunque coppia di nodi;
3. un grafo non orientato connesso, con n nodi e $n - 1$ lati.

1.1 Equivalenza delle definizioni

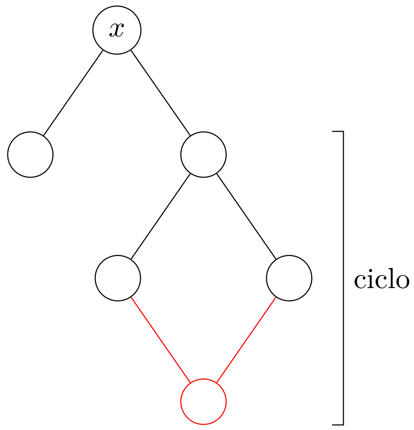
La 2 è equivalente alla 1 perché, se esistessero due cammini diversi tra una coppia di nodi, formerebbero un ciclo:



La 3 è equivalente alla 1 perché, scelto un nodo x qualsiasi dell'albero, tutti gli altri devono essere connessi a esso (essendo l'albero un grafo connesso). Per connettere gli altri $n-1$ nodi a x, servono almeno $n-1$ lati, e se ce ne sono di più esistono dei cicli.

Si può infatti immaginare di "sollevare" l'albero dal nodo x: allora, verrebbero sollevati anche i nodi adiacenti a x, che vi sono "attaccati" tramite un lato, poi i nodi a loro volta adiacenti a questi, e così via. Per poter sollevare tutti i nodi (cioè fare sì che il grafo sia connesso), ognuno degli $n - 1$ lati dovrebbe "tenere su" un nodo diverso. Se invece ci fossero più di $n - 1$ lati, esisterebbe almeno un nodo che viene sollevato da due o più lati, e allora si formerebbe un ciclo.

Ad esempio, questo grafo connesso non è un albero perché ha $n = 6$ nodi e $6 > n - 1$ lati, quindi deve contenere almeno un nodo sollevato da due lati, e di conseguenza un ciclo:



2 Albero con radice

Un **albero con radice** è una coppia $\langle r, T \rangle$ dove

- T è un albero;
- r è un nodo di T, che prende il nome di **radice**.

3 Livello

Il **livello** o **profondità** di un nodo è la lunghezza del cammino (semplice) che collega la radice al nodo.

4 Relazioni tra nodi

Dati due nodi x, y di un albero (con radice),

- x è **padre** di y, e y è **figlio** di x, se x precede immediatamente y nel cammino semplice dalla radice a y;
- x è **fratello** di y se x e y hanno lo stesso padre;
- x è **antenato** o **avo** di y, e y è **discendente** di x, se x compare nel cammino semplice dalla radice a y;

A volte si usano anche i termini **nonno** (padre del padre) e **nipote** (figlio di un figlio).

5 Nodi interni e foglie

In un albero, i nodi si dividono in

nodi interni: hanno figli;

foglie (o nodi esterni): non hanno figli.

6 Grado

Il **grado** di un nodo è il numero di figli posseduti.

Un albero si dice k**-ario** se il nodo con più figli ha grado k, cioè se

- i nodi hanno al massimo grado k;
- esiste almeno un nodo di grado k.

7 Numero di nodi per livello

In un albero k-ario ci sono al massimo k^i nodi di livello i.

Un livello si dice **saturo** se ha il massimo numero possibile di nodi.

8 Altezza

Un albero con radice $\langle r, T \rangle$ si può trasformare in un grafo orientato, orientando ogni lato $\{x, y\}$ dal padre verso il figlio.

Si dicono allora

- **altezza di un nodo** x la lunghezza del più lungo cammino da x a una foglia (nel sottoalbero con radice x);
- **altezza dell'albero** l'altezza della radice r.

9 Lunghezza del cammino

- La **lunghezza del cammino interno** LC_i di un albero è la somma dei livelli di tutti i *nodi interni*.
- La **lunghezza del cammino esterno** LC_e di un albero è la somma dei livelli di tutte le *foglie*.
- La **lunghezza del cammino** LC di un albero è la somma dei livelli di tutti i nodi, cioè la somma della lunghezza del cammino interno e della lunghezza del cammino esterno:

$$LC = LC_i + LC_e$$

10 Albero con radice ordinato

La relazione "essere fratello di" è una relazione di equivalenza sull'insieme dei nodi di un albero: le classi di equivalenza sono insiemi massimali di fratelli.

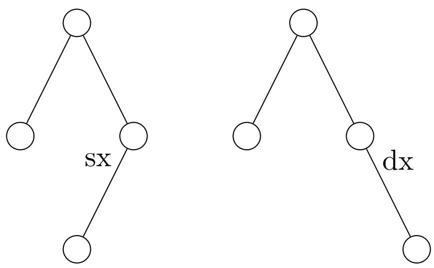
Un **albero con radice ordinato** è un albero con radice in cui ogni classe di equivalenza rispetto alla relazione "essere fratello di" è totalmente ordinata.

11 Albero binario

Un **albero binario** è un albero con radice ordinato in cui

- i nodi hanno al massimo due figli (ovvero hanno al massimo grado 2);
- si distinguono **figlio sinistro** e **figlio destro** di un nodo.

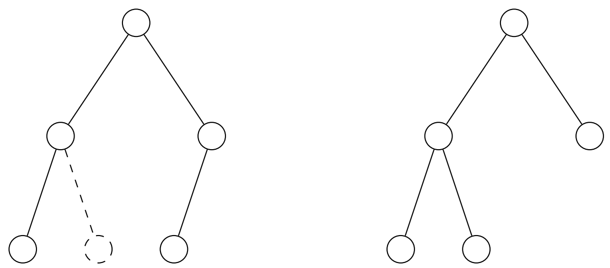
Alberi binari diversi possono corrispondere a uno stesso albero con radice ordinato, poiché in quest'ultimo non si fa distinzione tra figlio sinistro e destro:



12 Albero binario pieno e completo

Un albero binario si dice

- **pieno** se tutti i livelli sono saturi, tranne eventualmente l'ultimo;
- **completo** se è pieno e i nodi sull'ultimo livello sono tutti disposti il più a sinistra possibile, senza "buchi".



Albero binario *pieno*   Albero binario *completo*

Osservazioni: Per n nodi,

- un albero binario pieno (o anche completo) ha l'altezza *minima* possibile;
- esistono diversi alberi binari pieni, ma un unico albero binario completo (se si considera solo la "forma", e non il valore dei nodi).

12.1 Altezza e numero di nodi

Un albero binario completo di altezza h ha un numero di nodi n che soddisfa la relazione

$$2^h \leq n \leq 2^{h+1} - 1$$

Infatti:

- l'ultimo livello, h, ha da 1 a 2^h nodi;
- i livelli saturi precedenti, da 0 a $h - 1$, hanno

  $$\sum_{i=0}^{h-1} 2^i = 2^h - 1$$

  nodi.

Quindi, il numero minimo di nodi è $2^h - 1 + 1 = 2^h$, e il massimo è $2^h - 1 + 2^h = 2^h \cdot 2 - 1 = 2^{h+1} - 1$.

Viceversa, l'altezza soddisfa la relazione $h \sim \log_2 n$.

Gli stessi calcoli possono essere effettuati per altri tipi di alberi. Ad esempio, in un albero ternario[1] completo:

- l'ultimo livello, h, ha da 1 a 3^h nodi;
- i livelli saturi precedenti, $0, \dots, h-1$, hanno un numero di nodi totale pari a

  $$\sum_{i=0}^{h-1} 3^i = \frac{3^h - 1}{2}$$

  perché

  $$\sum_{i=0}^{h-1} 3^i = \underbrace{1111\dots1}_{h}{}_3$$
  $$2 \cdot \underbrace{1111\dots1}_{h}{}_3 + 1 = 1\underbrace{0000\dots0}_{h}{}_3 = 3^h$$
  $$\underbrace{1111\dots1}_{h}{}_3 = \frac{3^h - 1}{2}$$

Allora

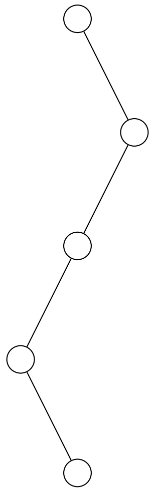
$$\frac{3^h - 1}{2} + 1 \le n \le \frac{3^h - 1}{2} + 3^h$$
$$\frac{3^h + 1}{2} \le n \le \frac{3^{h+1} - 1}{2}$$

e, viceversa, $h \sim \log_3 n$.

[1] Definito analogamente agli alberi binari, ovvero come un albero nel quale ogni nodo può avere al massimo tre figli, distinti tra sinistro, centrale e destro.

13 Albero binario degenere

Un albero binario si dice **degenere** se ha n nodi e altezza $n-1$, cioè se ogni nodo interno ha un solo figlio.



Osservazioni: Per n nodi,

- un albero binario degenere ha l'altezza *massima* possibile;
- esistono (considerando solo la "forma") 2^{n-1} alberi binari degeneri, dato che essi hanno una corrispondenza biunivoca con le stringhe binarie di lunghezza $n-1$: per ogni nodo, eccetto la radice, bisogna scegliere se posizionarlo come figlio sinistro o destro[2].

14 Albero binario localmente completo

Un albero binario si dice **localmente completo** se ogni nodo ha 0 o 2 figli.

Note:

- per alcuni autori, questa è la definizione di albero binario;
- nell'implementazione, un albero binario qualsiasi si può rendere localmente completo aggiungendo un "figlio fittizio", condiviso tra tutti i nodi che non hanno figli (effettivamente, questo nodo è l'unica foglia): ciò consente, in alcuni casi, di scrivere codice più efficiente, poiché non è necessario controllare se ciascuno dei due figli è presente o meno.

[2] Analogamente, esistono 3^{n-1} alberi ternari degeneri.

14.1 Numero di foglie

Teorema: Un albero binario localmente completo con n nodi interni ha $n+1$ foglie.

Dimostrazione: Si effettua per induzione.

- *Caso base*: Con 0 nodi interni, c'è 1 foglia, la radice.
- *Passo di induzione*: Tolta la radice, rimangono $n-1$ nodi interni, di cui $k-1$ nel sottoalbero sinistro e $n-k$ in quello destro. Per ipotesi di induzione, allora, il numero totale di foglie è

$$\underbrace{((k-1)+1)}_{\text{sinistra}} + \underbrace{((n-k)+1)}_{\text{destra}} = n+1 \qquad \square$$

La condizione che l'albero sia localmente completo serve a garantire l'esistenza sia del sottoalbero sinistro che di quello destro per tutti i nodi interni, in modo da poter applicare l'ipotesi di induzione.

14.2 Numeri di Catalano

Teorema: Il numero di alberi binari localmente completi con n nodi interni è

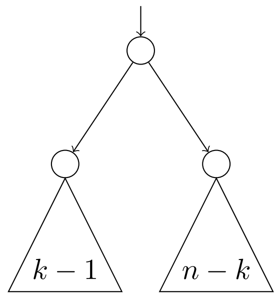
$$\frac{1}{n+1}\binom{2n}{n}$$

Questa formula è la soluzione dell'equazione di ricorrenza

$$\begin{aligned} c_0 &= 1 \\ c_n &= \sum_{k=1}^{n} c_{k-1}c_{n-k} \quad n > 0 \end{aligned}$$

che definisce la successione dei **numeri di Catalano**, corrispondenti al numero di alberi binari localmente completi.

Infatti, con 0 nodi interni, c'è solo un albero, formato unicamente dalla radice. Con $n > 0$ nodi interni, invece, la radice ha 2 figli (perché l'albero è localmente completo). Ciascuno di essi è a sua volta radice di un sottoalbero localmente completo, e i due sottoalberi devono contenere in totale $n-1$ nodi interni. Se il sottoalbero sinistro ha $k-1$ nodi interni, allora quello destro ne contiene $n-k$.



Di conseguenza, per ogni $k \in \mathbb{N}$, $1 \leq k \leq n$, ci sono c_{k-1} sottoalberi sinistri e c_{n-k} sottoalberi destri tra cui scegliere: in totale $c_{k-1}c_{n-k}$ scelte. c_n è quindi la somma del numero di scelte per ogni valore di k.

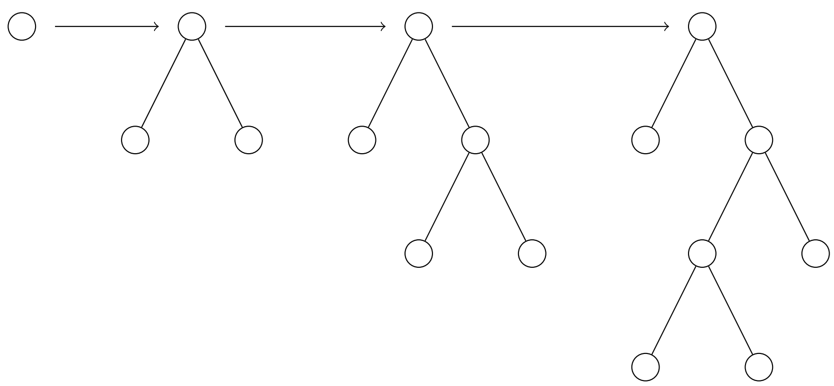
14.3 Lunghezza del cammino

Teorema: In un albero binario localmente completo con n nodi interni, vale la relazione

$$LC_e = LC_i + 2n$$

14.3.1 Dimostrazione

Ogni albero binario localmente completo si può ottenere partendo dalla sola radice, e trasformando una foglia alla volta in un nodo interno. Ad esempio:



- L'albero di partenza, formato dalla sola radice, ha $LC_i = LC_e = 0$.
- Quando una foglia al livello k si trasforma in nodo interno, si elimina il suo contributo a LC_e, pari a k, che viene invece aggiunto a LC_i, ma a LC_e si sommano i livelli $k+1$ delle due nuove foglie:

$$LC_i := LC_i + k$$
$$LC_e := LC_e - k + 2(k+1) = LC_e + k + 2$$

Siccome la differenza di incremento è 2, dopo n passi, ovvero dopo la creazione di n nodi interni, si ha

$$LC_e = LC_i + 2n \qquad \square$$

14.3.2 Dimostrazione per induzione

- *Caso base*: L'albero con $n = 0$ nodi interni, cioè formato dalla sola radice, ha

$$LC_e = LC_i = 0$$
$$LC_e = LC_i + 2 \cdot 0$$

- *Passo di induzione*: Tolta la radice, rimangono due sottoalberi:
  - il sottoalbero sinistro, con $k - 1$ nodi interni e lunghezze dei cammini LC_i^{sx} e LC_e^{sx};
  - il sottoalbero destro, con $n - k$ nodi interni e lunghezze dei cammini LC_i^{dx} e LC_e^{dx}.

  La radice si trova al livello 0, quindi non contribuisce direttamente alle lunghezze dei cammini. Rimuovendola, però, il livello di tutti gli altri nodi diminuisce di 1. Di conseguenza, la lunghezza totale del cammino interno dei due sottoalberi diminuisce di $n-1$ (il numero di nodi interni, meno la radice), mentre la lunghezza del cammino esterno si riduce di $n+1$ (il numero di foglie, invariato perché la radice rimossa era un nodo interno):

$$LC_i^{sx} + LC_i^{dx} = LC_i - (n - 1)$$
$$LC_e^{sx} + LC_e^{dx} = LC_e - (n + 1)$$

  Per ipotesi di induzione, nei sottoalberi valgono le relazioni

$$LC_e^{sx} = LC_i^{sx} + 2(k - 1) = LC_i^{sx} + 2k - 2$$
$$LC_e^{dx} = LC_i^{dx} + 2(n - k) = LC_i^{dx} + 2n - 2k$$

e allora, complessivamente,

$$\begin{aligned}
LC_e^{sx} + LC_e^{dx} &= LC_i^{sx} + 2k - 2 + LC_i^{dx} + 2n - 2k \\
\underbrace{LC_e^{sx} + LC_e^{dx}}_{LC_e-(n+1)} &= \underbrace{LC_i^{sx} + LC_i^{dx}}_{LC_i-(n-1)} + 2n - 2 \\
LC_e - n - 1 &= LC_i - n + 1 + 2n - 2 \\
LC_e &= LC_i + 2n + 2 - 2 \\
LC_e &= LC_i + 2n \qquad \square
\end{aligned}$$

15 Albero di copertura

Dato un grafo $G = \langle V, E \rangle$, non orientato e connesso, un **albero di copertura** di G è un sottografo $T = \langle V, E' \rangle$, $E' \subseteq E$ che

- ha tutti i vertici di G;
- è un albero.