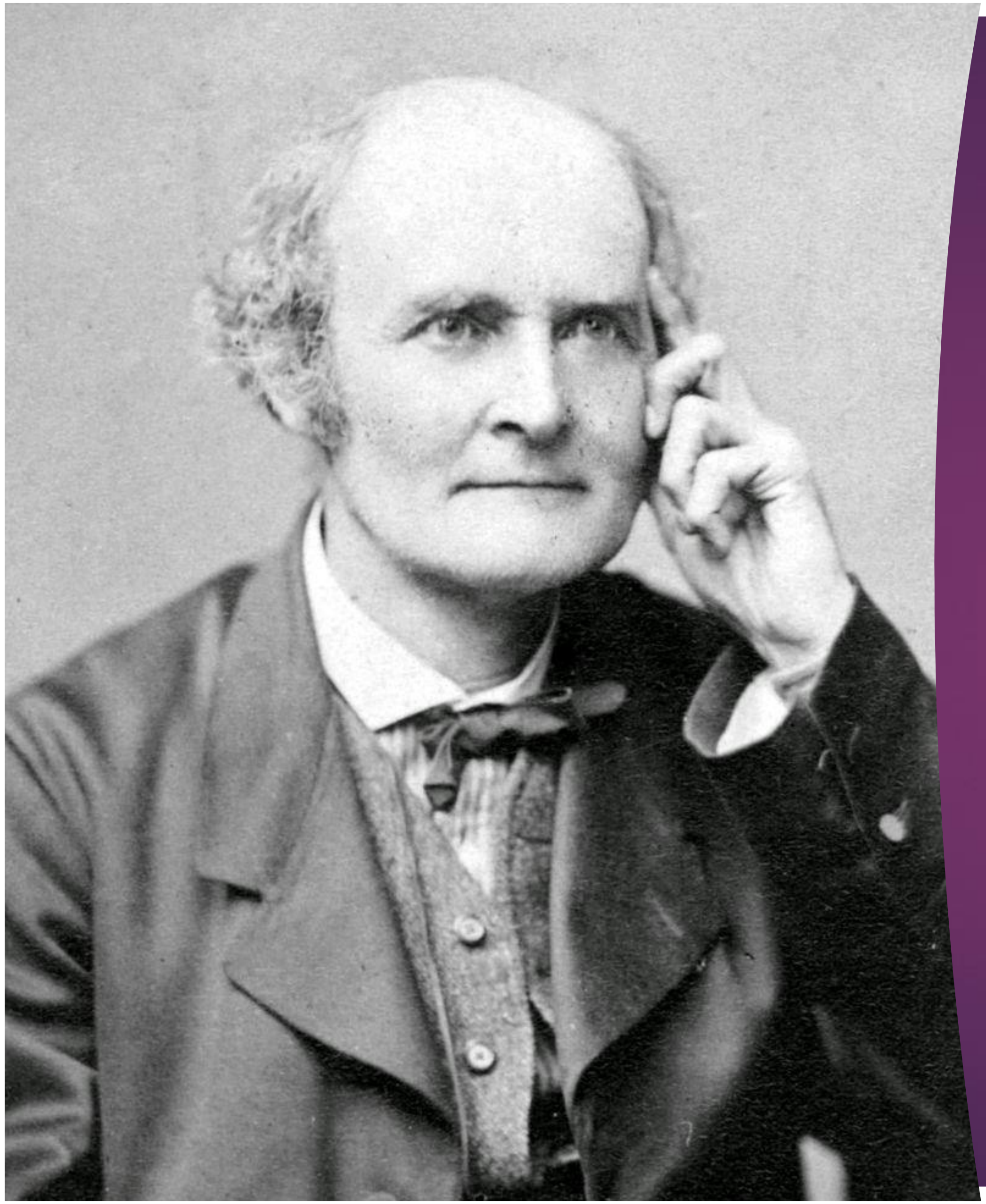


La Formula di Cayley

SEMINARIO DI TEORIA DEI GRAFI

Argomenti trattati

- Introduzione ai grafi orientati
- La Formula di Cayley
- Applicazioni del teorema

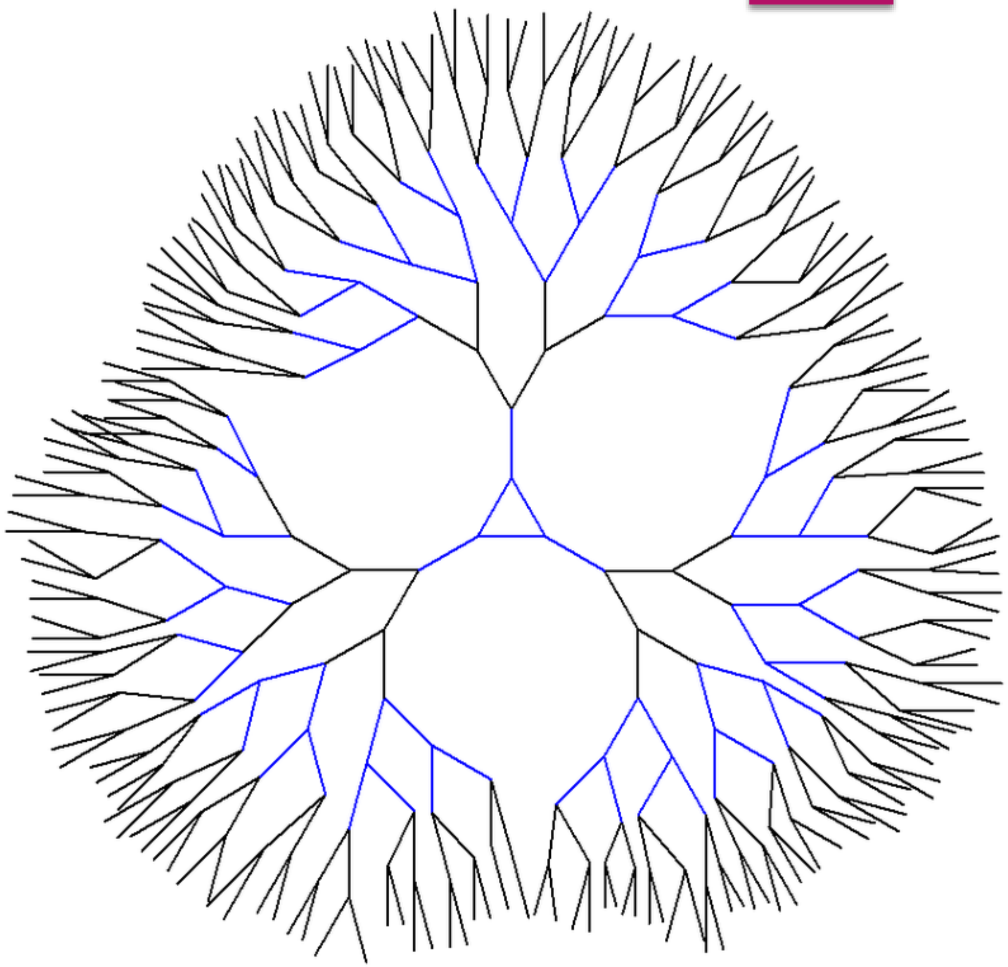
Anteprima


La Formula di Cayley
SEMINARIO DI TEORIA DEI GRAFI
Anteprima
Cenno storico e introduzione al problema
Matematico inglese, nato a Richmond (Surrey) il 16 agosto 1821, morto a Cambridge il 26 gennaio 1895. Esercitò a Londra fino a 42 anni la professione legale, pur non interrompendo mai l'intensa produzione matematica iniziata nel 1841; e nel 1863 fu chiamato alla cattedra di algebra, detta sadleriana, dell'università di Cambridge, che tenne fino alla morte.
Fonti accertate riferiscono che fu Cayley uno dei primi ad introdurre nella chimica molecolare l'utilizzo di quegli oggetti che oggi noi chiamiamo grafi - in realtà in una sua pubblicazione usò questi oggetti per approfondire la struttura molecolare di alcuni isomeri, ancora prima di coniare il termine 'grafo'. Dunque la celebre 'Formula di Cayley' deriva dagli studi del matematico riguardo la 'Teoria Chimica dei Grafi'.
Grafo molecolare che descrive la struttura di un dendrimero.
Introduzione ai grafi etichettati
Grafi orientati
Alberi radicati
La Formula di Cayley
Applicazioni
Referenze
M. Melo, lezioni di GE460 - Teoria dei Grafi, Università di Roma Tre, anno 2019/2020.
J. A. Bondy – U. S. R. Murty, Graph Theory, Springer, 2008.
Grazie per l'attenzione

Cenno storico e introduzione al problema

Matematico inglese, nato a Richmond (Surrey) il 16 agosto 1821, morto a Cambridge il 26 gennaio 1895. Esercitò a Londra fino a 42 anni la professione legale, pur non interrompendo mai l'intensa produzione matematica iniziata nel 1841; e nel 1863 fu chiamato alla cattedra di algebra, detta sadleriana, dell'università di Cambridge, che tenne fino alla morte.

Fonti accertate riferiscono che fu Cayley uno dei primi ad introdurre nella chimica molecolare l'utilizzo di quegli oggetti che oggi noi chiamiamo grafi - in realtà in una sua pubblicazione usò questi oggetti per approfondire la struttura molecolare di alcuni isomeri, ancora prima di coniare il termine 'grafo'. Dunque la celebre 'Formula di Cayley' deriva dagli studi del matematico riguardo la 'Teoria Chimica dei Grafi'.



Grafo molecolare che descrive la struttura di un dendrimero.

Introduzione ai grafi etichettati

Introduzione

La Formula di Cayley - che tratteremo più avanti – ha bisogno di un'introduzione ai grafi orientati, poiché essi ci permettono con maggiore agilità di dimostrarlo. All'inizio del corso abbiamo introdotto i grafi orientati, senza però approfondire alcuni loro aspetti. Inoltre l'enunciato del teorema parla di grafi etichettati (**labelled graph**); quindi la prima cosa che definiremo sono i grafi etichettati, anche grazie ad esempi.

Alcune definizioni verranno date per assimilate, mentre altre verranno inserite all'interno di tale seminario come complemento delle lezioni fino ad ora fatte.

Grafi etichettati

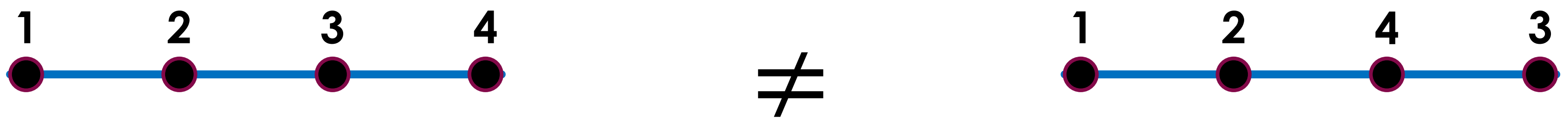
Definizione: Un grafo G si dice **etichettato** se i vertici sono muniti di una etichetta.

Nel corso ci siamo imbattuti in moltissimi esempi di grafi etichettati. Il fatto stesso di associare ai vertici delle lettere o numeri significa munirli di etichetta.

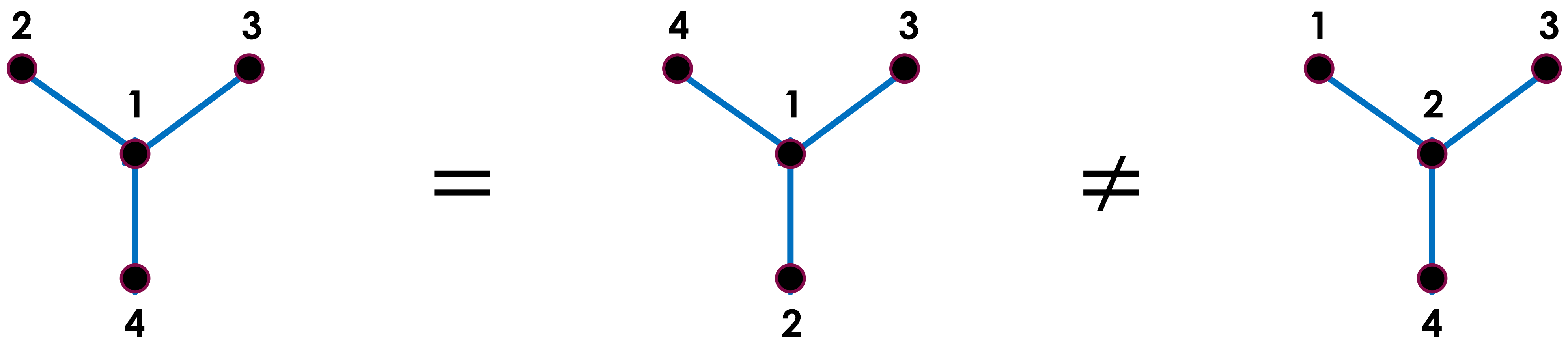
Ad esempio nel seminario della collega Barbara Galati, alcune dimostrazioni sono state fatte con l'uso di un insieme etichettato, quale l'insieme dei colori. Dobbiamo stare attenti quando usiamo le etichette perché grafi che sembrano apparentemente uguali, in realtà non lo sono.

Esempi: Quanti alberi etichettati distinti ci sono con 4 vertici?

1. Grafi semplici che sono cammini in 4 vertici possono non risultare essere uguali. Sia l'insieme di etichette $\boldsymbol{M} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}$. Allora i due cammini disegnati sotto non sono uguali poiché, ad esempio, i vertici non rispettano i gradi. In realtà basta imporre che i vertici rispettino la proprietà di adiacenza per avere un isomorfismo.



1. Un altro esempio di grafo semplice in 4 vertici, diverso dall'esempio precedente, è la stella a tre punte con un vertice centrale di grado 3.



In realtà, se consideriamo il numero di vertici $|\mathbf{V}(\boldsymbol{G})| = \mathbf{4}$, gli unici tipi di grafi che possiamo ottenere sono quelli mostrati nei due esempi precedenti. Inoltre, con molta facilità, possiamo contare il numero di grafi etichettati distinti a due a due in entrambi i casi.

- Nel caso dell'**Esempio 2.**, ho in tutto 4 stelle distinte, poiché l'unico vertice che determina un grafo etichettato diverso dagli altri è quello centrale; fissato il vertice etichettato centrale, allora ho 3!=6 grafi etichettati isomorfi.

- Nel caso dell' **Esempio 1.**, quindi di cammini con 4 vertici, basta contare le possibilità di etichettatura dei 4 vertici che sono 4!, ma stare attenti che i grafi presi in considerazione sono uguali a grafi con etichetta iniziale e finale invertiti. Quindi ho

$$\frac{\mathbf{4!}}{\mathbf{2}} = \mathbf{12}$$

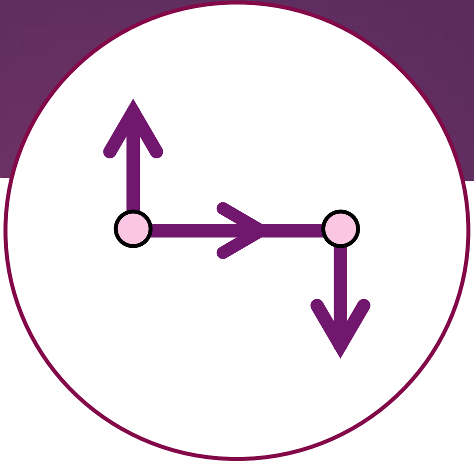
cammini etichettati distinti con 4 vertici. Quindi in tutto abbiamo 16 grafi etichettati distinti in 4 vertici.

Ma cosa succede se prendo un grafo etichettato con n vertici? Questo ragionamento non può essere applicato per n molto grandi, perché avrei moltissime possibilità di disporre e collegare i vertici. La Formula di Cayley risolve questo problema senza mettere alcuno sforzo.

Osservazione

Si potrà notare come ci sia una relazione che intercorre tra alberi etichettati in n vertici e alberi generanti di $\boldsymbol{K_n}$.

Nel prossimo paragrafo richiameremo alcune definizioni di oggetti matematici che sono il fulcro di questo seminario, nonché parte integrante della dimostrazione della Formula di Cayley.



Grafi orientati

Grafi orientati

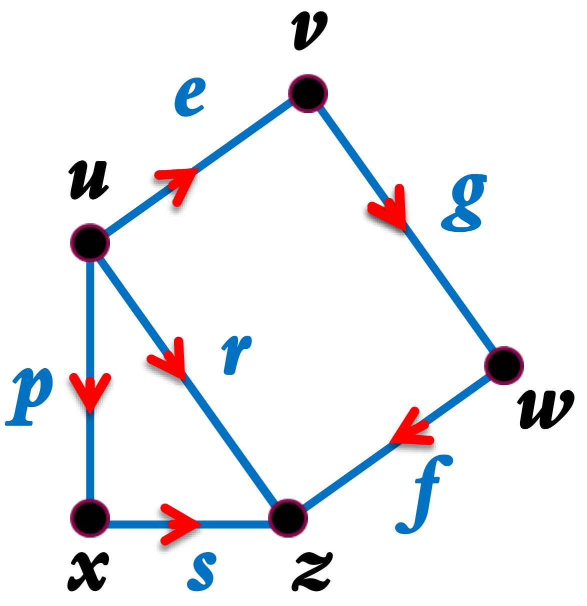
Definizione: Un grafo **orientato** è una terna $\boldsymbol{D} = (\boldsymbol{V}(\boldsymbol{D}), \boldsymbol{A}(\boldsymbol{D}), \boldsymbol{\psi_D})$ tale che

- $\boldsymbol{V}(\boldsymbol{D})$ è l'insieme dei vertici
- $\boldsymbol{A}(\boldsymbol{D})$ è l'insieme degli archi
- $\boldsymbol{\psi_d}(\boldsymbol{a}) = (\boldsymbol{u}, \boldsymbol{v})$ è una coppia **ordinata** di vertici

In un certo senso quindi possiamo pensare i grafi orientati come un grafo con l'aggiunta di un'orientazione degli archi, cioè con un ordine per ogni arco $\boldsymbol{a} \in \boldsymbol{V}(\boldsymbol{D})$.

La prima osservazione è che dato un grafo orientato $\boldsymbol{D}$, se non teniamo conto dell'orientazione, ciò che otteniamo è un sottografo di $\boldsymbol{D}$. Quindi dato un grafo $\boldsymbol{G}$, esistono diversi grafi orientati che hanno $\boldsymbol{G}$ come grafo sottostante. Ad esempio ogni grafo incontrato a lezione possiamo munirlo di un'orientazione in modo tale che il grafo sottostante sia $\boldsymbol{G}$.

Esempio di grafo orientato



Usando le notazioni appena introdotte, scriveremo:

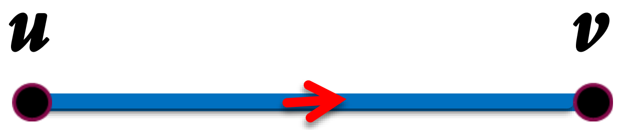
$$\psi_D(e) = (u, v) \qquad \psi_D(r) = (u, z)$$

$$\psi_D(g) = (v, w) \qquad \psi_D(s) = (s, z)$$

$$\psi_D(f) = (w, z) \qquad \psi_D(p) = (u, x)$$

Notazioni e definizioni

- **Definizione:** Dato un arco $\boldsymbol{a} \in \boldsymbol{A}(\boldsymbol{D})$, se $\boldsymbol{a} = (\boldsymbol{u}, \boldsymbol{v})$, allora:
  - $\boldsymbol{u}$ si dice **coda** di $\boldsymbol{a}$
  - $\boldsymbol{v}$ si dice **testa** di $\boldsymbol{a}$



- **Definizione:** Dato un vertice $\boldsymbol{v} \in \boldsymbol{V}(\boldsymbol{D})$, allora definiamo:
  - $\boldsymbol{indeg}(\boldsymbol{v}) = \boldsymbol{d}_G^-(\boldsymbol{v}) =$ il numero di archi che hanno testa in $\boldsymbol{v}$.
  - $\boldsymbol{outdeg}(\boldsymbol{v}) = \boldsymbol{d}_G^+(\boldsymbol{v}) =$ il numero di archi che hanno coda in $\boldsymbol{v}$.

Alcune osservazioni

- **Osservazione 1**: Dato $\boldsymbol{D}$ grafo orientato e $\boldsymbol{G}$ il suo grafo sottostante, allora per ogni $\boldsymbol{v} \in \boldsymbol{V}(\boldsymbol{G})$ vale seguente relazione tra i gradi:

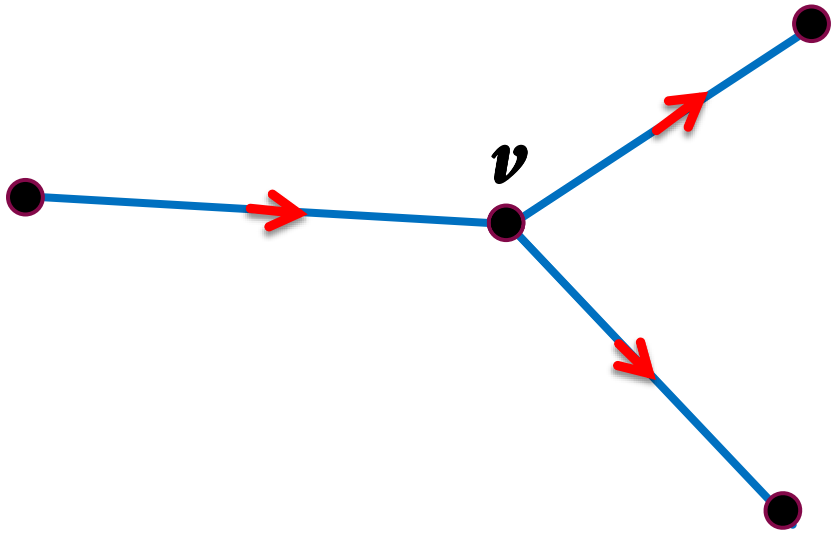
$$\boldsymbol{d_G(v) = d_G^-(v) + d_G^+(v)}$$

- **Osservazione 2**: Vale una formula analoga all'**Handshaking Lemma**, i.e.

$$\boldsymbol{\sum_{v \in V} d_D^-(v) = |A(D)|}$$

- **Osservazione 3**: In un grafo orientato ha senso parlare di cammini diretti, anziché di solamente cammini.

Illustrazione 1



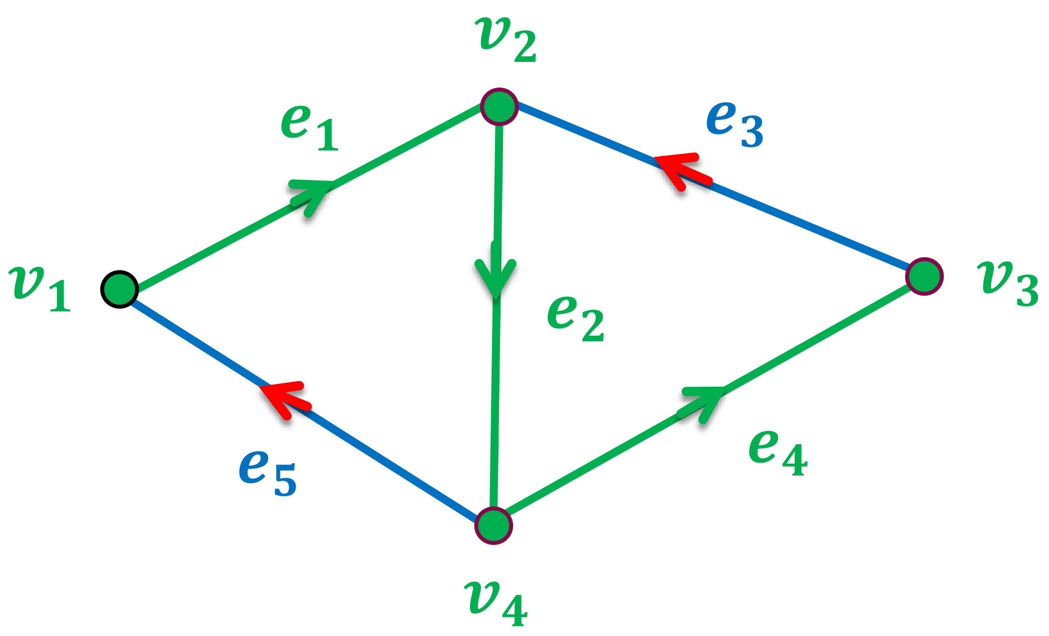
In questo esempio, grazie alle notazioni introdotte precedentemente, possiamo scrivere:

$$d_G^-(v) = 1$$

$$d_G^+(v) = 2$$

$$\sum_{v \in V} d_D^-(v) = |A(D)| = 3$$

Illustrazione 2



In questa illustrazione, prendendo in considerazione il cammino diretto

$$v_1 e_1 v_2 e_2 v_4 e_4 v_2$$

nel grafo $\boldsymbol{D}$ orientato tra $\boldsymbol{v_1}$ e $\boldsymbol{v_3}$.

Definizione: Se esiste un cammino diretto tra due vertici $\boldsymbol{u}$ e $\boldsymbol{v}$, allora diremo che $\boldsymbol{v}$ è **raggiungibile** da $\boldsymbol{u}$.

Altre definizioni

Definizione: Sia $\boldsymbol{D}$ un grafo orientato. Se esiste $\boldsymbol{v} \in \boldsymbol{V}(\boldsymbol{D})$ tale che

$$\boldsymbol{indeg}(\boldsymbol{v}) = \mathbf{0} \quad (\Leftrightarrow \boldsymbol{outdeg}(\boldsymbol{v}) = \boldsymbol{d_G}(\boldsymbol{v}))$$

allora $\boldsymbol{v}$ è detta **sorgente** (**source**).

Definizione: Sia $\boldsymbol{D}$ un grafo orientato. Se esiste $\boldsymbol{v} \in \boldsymbol{V}(\boldsymbol{D})$ tale che

$$\boldsymbol{outdeg}(\boldsymbol{v}) = \mathbf{0} \quad (\Leftrightarrow \boldsymbol{indeg}(\boldsymbol{v}) = \boldsymbol{d_G}(\boldsymbol{v}))$$

allora $\boldsymbol{v}$ è detta **lavello** (**sink**).

Alberi radicati

Definizione e uso degli alberi radicati

- **Definizione:** Un albero **radicato** è un albero $\boldsymbol{T}$ con una scelta di un suo vertice $\boldsymbol{x} \in \boldsymbol{V}(\boldsymbol{T})$, chiamato **radice** di $\boldsymbol{T}$.

- **Osservazione:** Dato un albero $\boldsymbol{T}$ è possibile orientare il grafo in modo tale che, scelta una radice $\boldsymbol{x} \in \boldsymbol{V}(\boldsymbol{T})$, vale che

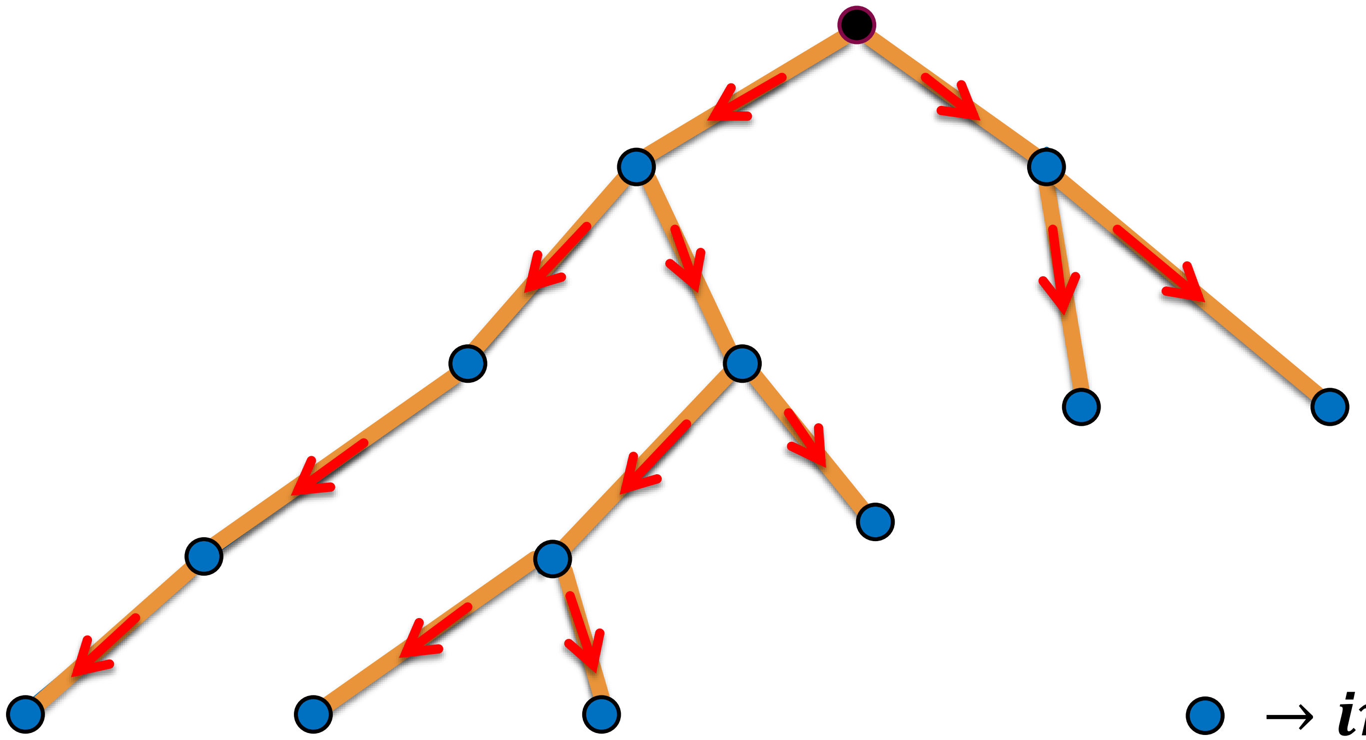
$$\boldsymbol{indeg}(\boldsymbol{v}) = \mathbf{1}, \qquad \forall\, \boldsymbol{v} \in \boldsymbol{V}(\boldsymbol{T}) \setminus \{\boldsymbol{x}\} \qquad (\cdot)$$

- **Definizione:** Se in un albero $\boldsymbol{T}$ scegliamo un'orientazione come definita nell'**Osservazione** precedente, allora tale orientazione viene chiamata **arborescenza** o $\boldsymbol{x}$ **– arborescenza** (**branching** o $\boldsymbol{x}$ **- branching**).

Da notare come tale orientazione rende sorgente solamente la radice $\boldsymbol{x}$ e che possiamo associare ad un albero radicato la $\boldsymbol{x}$ – arborescenza. Viceversa, data una $\boldsymbol{x}$ – arborescenza di un albero $\boldsymbol{T}$, possiamo in modo analogo associare ad essa un unico albero radicato in $\boldsymbol{x}$. Questa osservazione (andrebbe dimostrata) fa vedere la corrispondenza biunivoca tra gli alberi radicati in $\boldsymbol{x}$ e le $\boldsymbol{x}$ – arborescenze.

Esempio di albero radicato con x - arborescenza

$x \rightarrow indeg(x) = 0$



$\rightarrow indeg(x) = 1$

La Formula di Cayley

Enunciato

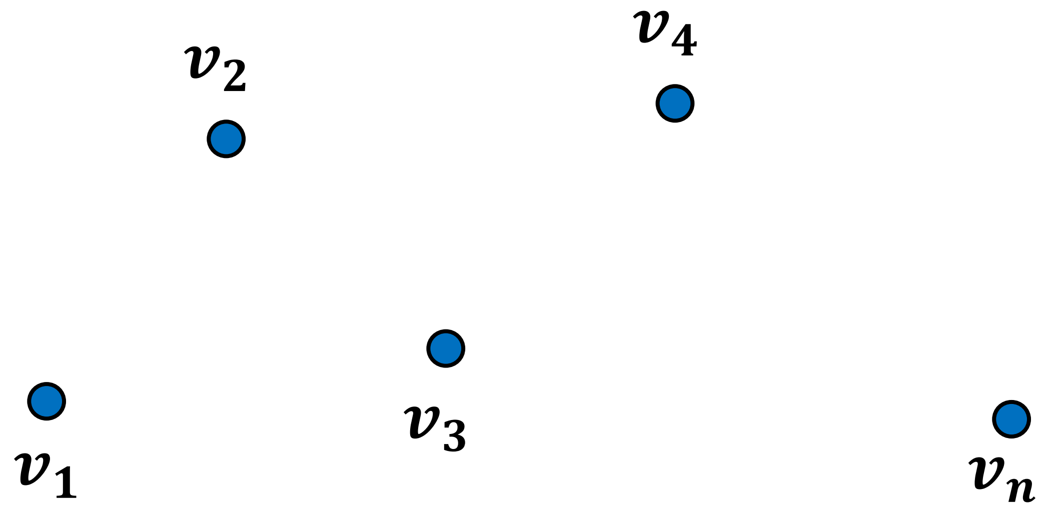
Teorema [Formula di Cayley]

Il numero di alberi etichettati in $\boldsymbol{n}$ vertici è $\boldsymbol{n^{n-2}}$.

Idea grafica

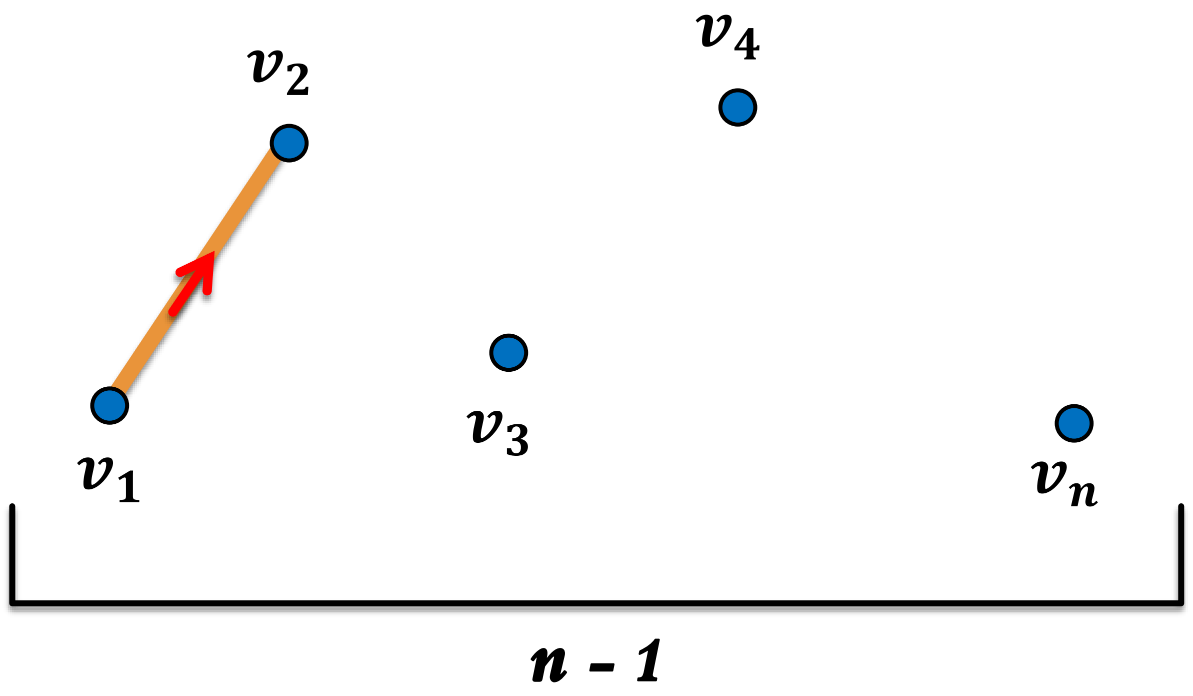
L'idea per dimostrare la Formula di Cayley è quella di contare il numero di arborescenze etichettate in n vertici e sapendo che un'arborescenza in un albero corrisponde alla scelta di un vertice, concludere che il numero di alberi etichettati è $\boldsymbol{n^{n-2}}$.

Passo 0: ho una foresta con $\boldsymbol{n}$ componenti connesse, date dagli $\boldsymbol{n}$ vertici isolati (**alberi**) che risultano essere tutte arborescenze.



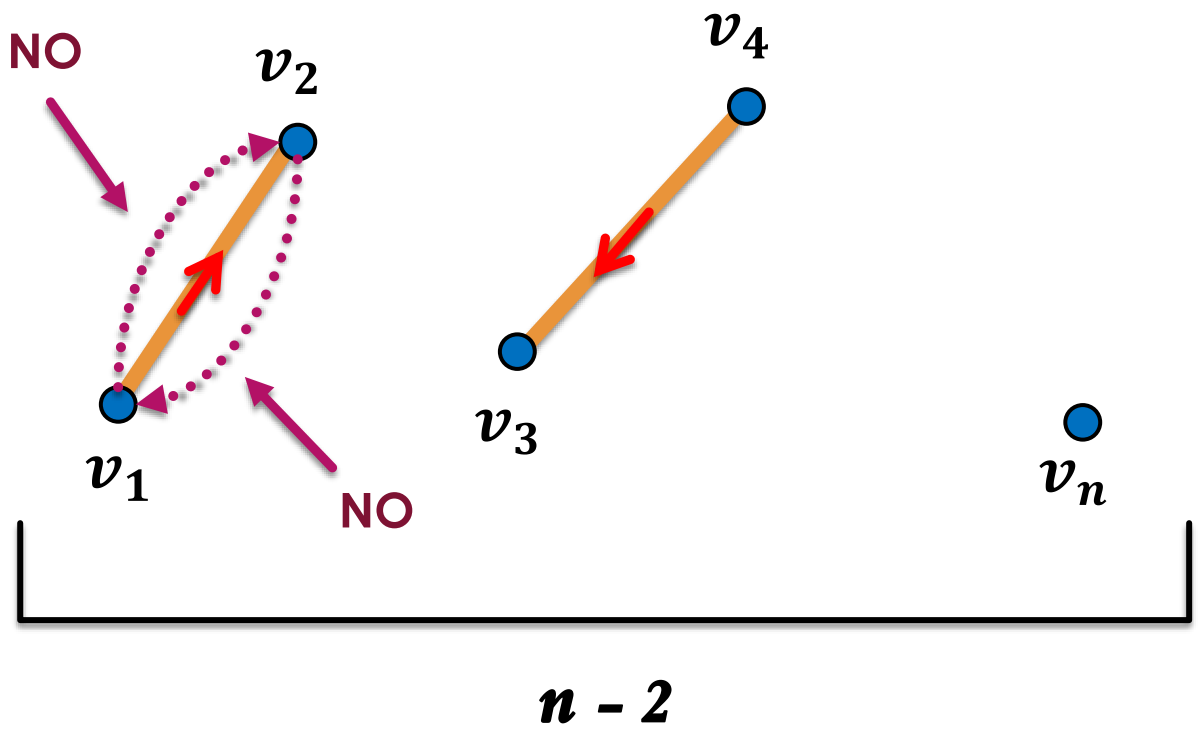
Idea grafica

Passo 1: Scegliamo un arco qualunque, scegliendo in modo arbitrario la testa e la coda. Ciò che ottengo è una foresta con $n - 1$ componenti connesse, ciascuna delle quali è una arborescenza.



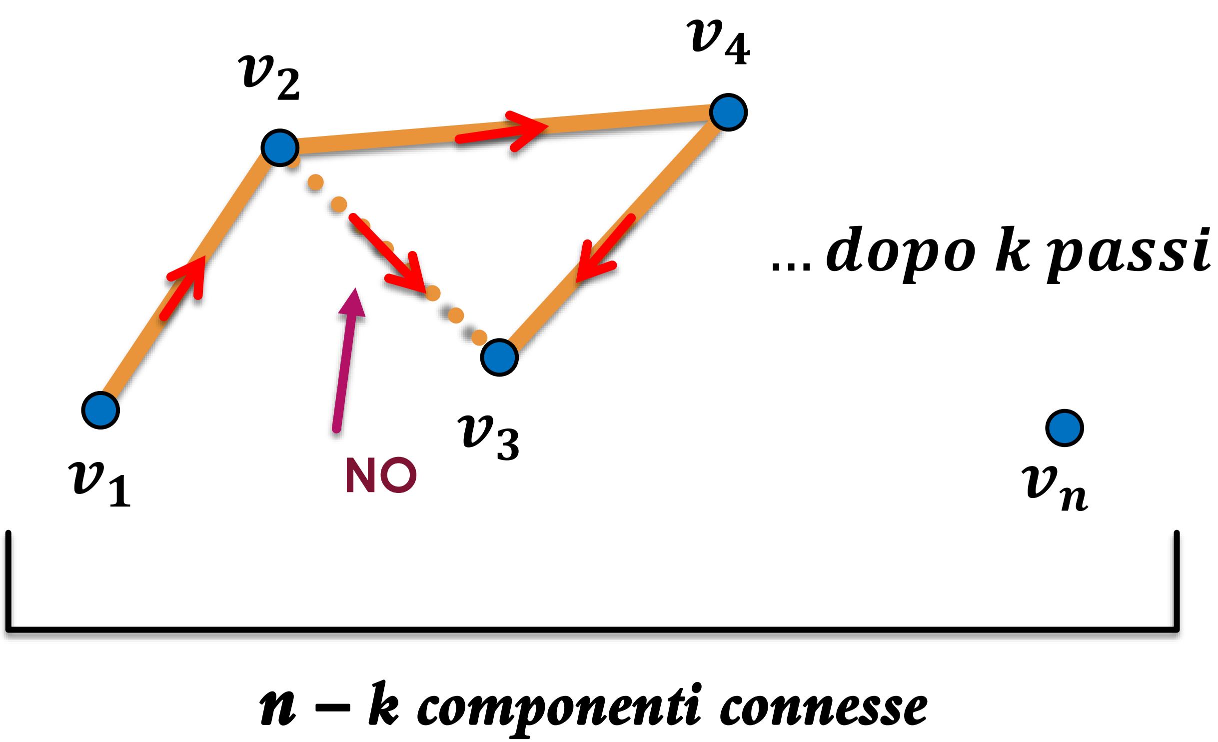
Idea grafica

Passo 2: Prendiamo un'altra coppia di vertici per formare un nuovo arco, ma questi due vertici devono essere in modo tale da non stare nella stessa componente connessa. In questo modo otteniamo $\boldsymbol{n-2}$ componenti connesse, ciascuna delle quali è un'arborescenza.



Idea grafica

Passo k: Nei successivi passi ciò che dobbiamo fare è creare un arco $(\boldsymbol{u}, \boldsymbol{v})$, dove u si può scegliere arbitrariamente tra gli $\boldsymbol{n}$ vertici, mentre $\boldsymbol{v}$ lo dobbiamo scegliere in modo che non sia in una componente connessa che contenga $\boldsymbol{u}$ e che sia sorgente per la componente connessa di arrivo, come mostrato nell'animazione. Iterando tale procedimento, al $\boldsymbol{k}$ – esimo passo, ottengo $\boldsymbol{n} - \boldsymbol{k}$ componenti connesse.



Dimostrazione

Cominciamo considerando $\boldsymbol{n}$ vertici isolati, quindi abbiamo una foresta composta da $\boldsymbol{n}$ componenti connesse.

La procedura che utilizziamo segue l'idea affrontata nella slide precedente.

Ad ogni passo, aggiungendo un lato, diminuiamo di uno il numero di componenti connesse.
Supponendo di aver costruito con tale procedimento $\boldsymbol{k}$ componenti connesse, il numero di scelte che abbiamo per costruire il nuovo arco $(\boldsymbol{u}, \boldsymbol{v})$ è $\boldsymbol{n}(\boldsymbol{k}-\mathbf{1})$, dove $\boldsymbol{n}$ è il numero di possibilità di scelta della coda $\boldsymbol{u}$, mentre $\boldsymbol{k}-\mathbf{1}$ è la limitazione sulle possibilità di scelta di $\boldsymbol{v}$ come testa del nuovo arco, il quale dovrà andare ad incidere su una sorgente di una componente connessa, fatta ad eccezione della componente connessa che contiene $\boldsymbol{u}$; infatti se per assurdo il nuovo arco avesse la testa $\boldsymbol{v}$ nella componente che contiene il vertice $\boldsymbol{u}$, si creerebbe un ciclo all'interno di un albero etichettato, il che è un assurdo.

In tale modo otteniamo il numero di modi possibili di costruire un'arborescenza al variare del numero dei lati che vado a tracciare. Quindi vale:

$$\prod_{i=1}^{n-1} n(n-i) = n^{n-1}(n-1)!$$

Questa produttoria però è fuorviante, poiché, quando passo dopo passo scelgo il lato da costruire - seguendo le proprietà illustrate nelle slide precedenti –, il risultato finale è la stessa arborescenza.
Segue che il numero eccedente di arborescenze che sto contando è $(n-1)!$ – l'ordine in cui è possibile scegliere una precisa orientazione.

In questo modo ciò a cui arriviamo è il numero di arborescenze etichettate A_e in n vertici è

$$|A_e| = \prod_{i=1}^{n-1} \frac{n(n-i)}{(n-1)!} = \frac{n^{n-1}\cancel{(n-1)!}}{\cancel{(n-1)!}} = n^{n-1}$$

Ma la costruzione di un'arborescenza è determinata dalla radice x che prendo in considerazione inizialmente e siccome sono n le scelte possibili delle radici, allora il numero che cerchiamo è:

$$|A_e| = n^{n-2}$$

■

Applicazioni

Prima applicazione della Formula di Cayley

- Nella sezione sui grafi etichettati c'eravamo posti un quesito:

« Quanti alberi etichettati distinti ci sono con 4 vertici? »

La risposta si può trovare senza fare tutti i possibili disegni poiché possiamo applicare la Formula di Cayley al problema ed ottenere la risposta:

$$n^{n-2} = 4^2 = 16$$

Inoltre se estendiamo questo quesito ai grafi etichettati in n vertici, si deduce velocemente che cercare il numero degli alberi generanti di K_n equivale a cercare il numero delle arborescenze distinte di K_n.

Si può concludere, quindi usando la Formula di Cayley.

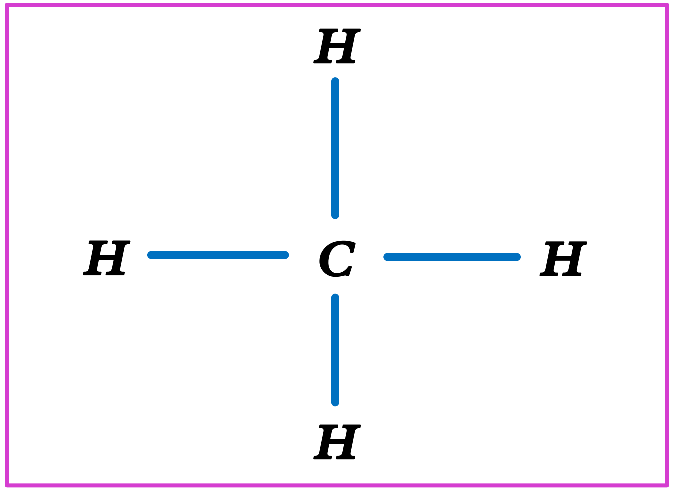
Seconda applicazione della Formula di Cayley

Nel 1857 Arthur Cayley – come già accennato – ha scoperto gli alberi, mentre sta enumerando tutti gli isomeri degli idrocarburi saturi, composti chimici che hanno forma $C_k H_{2k+2}$. Il suo lavoro fu importantissimo perché molti problemi di chimica e biochimica vengono risolti sfruttando i suoi lavori sviluppati in Teoria dei Grafi.

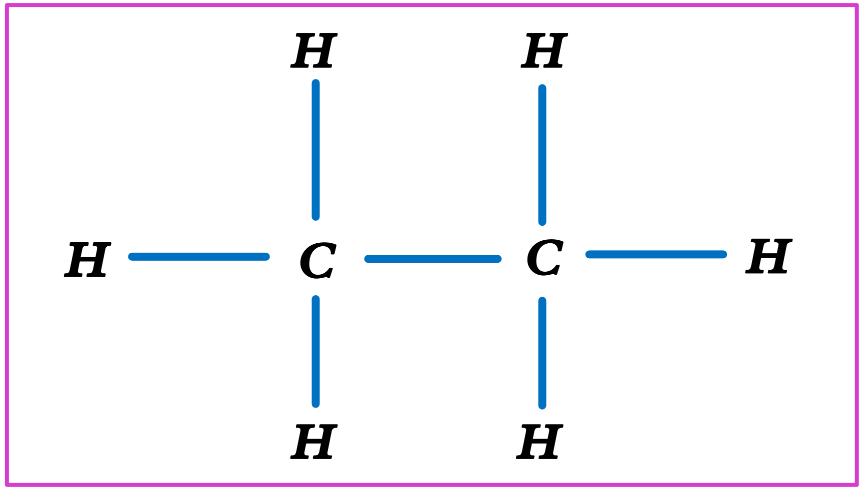
Possiamo rappresentare questi tipi di isomeri nel modo seguente: indichiamo con C ogni atomo di carbonio e con H gli atomi di idrogeno e tracciamo uno spigolo fra due di questi atomi se essi hanno un legame nel dato composto. I grafi ottenuti in questo modo sono noti come i **grafi dei legami chimici**.

Per esempio, qui di seguito sono rappresentati il metano e l'etano, dove C ha grado 4 e H ha grado 1.

Metano



Etano



Si osservi che le figure rappresentate nella slide precedente sono tutti alberi. Si dimostra che questo fatto non è una semplice coincidenza; infatti il grafo dei legami chimici di ogni idrocarburo saturo è un albero. Per dimostrare ciò indichiamo con $G = (V, E)$ il grafo dei legami chimici di $\boldsymbol{C_k H_{2k+2}}$.

Si ha che

$$|V| = k + 2k + 2 = 3k + 2$$

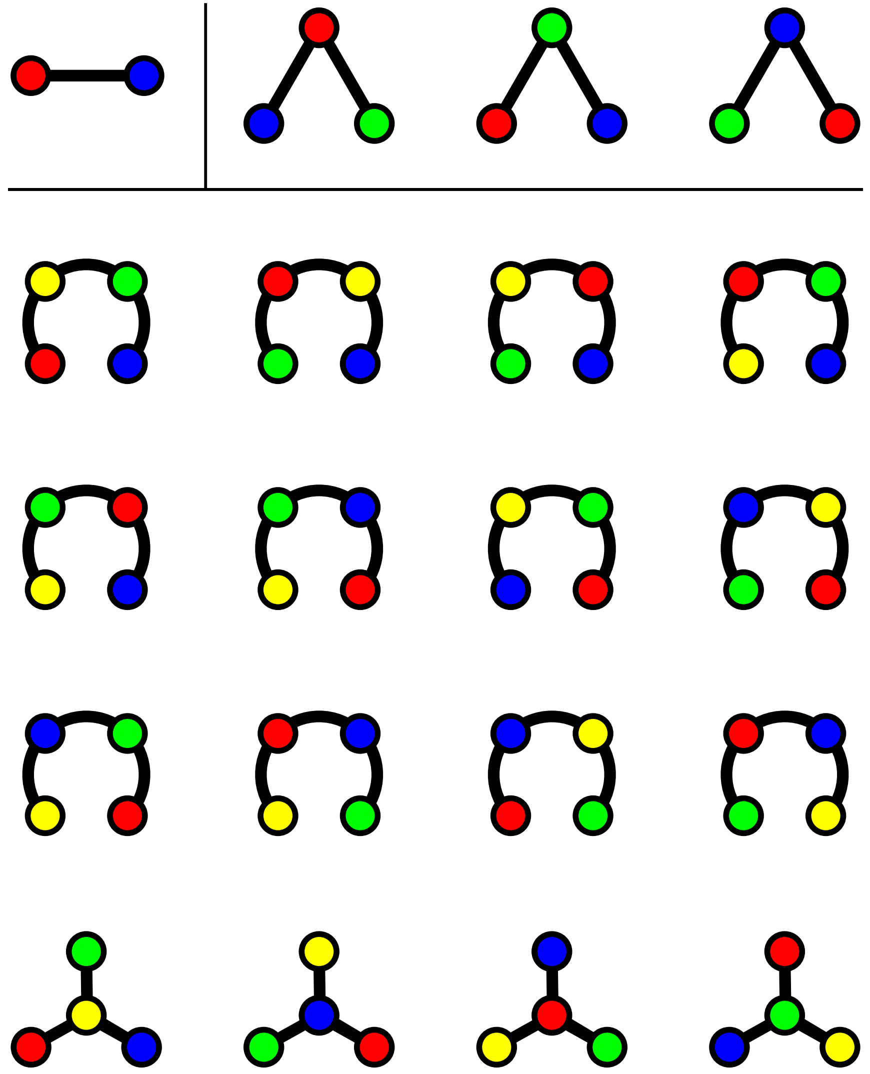
Ricordando che ogni atomo di carbonio ha grado 4 e ogni atomo di idrogeno ha grado 1, vale

$$\sum_{x \in V} d(x) = 4k + 1(2k + 2) = 6k + 2$$

e per l'Handshaking Lemma

$$|E| = \frac{\sum_{x \in V} d(x)}{2} = 3k + 1$$

Quindi $|E| = |V| - 1$. Poiché il grafo dei legami chimici di $\boldsymbol{C_k H_{2k+2}}$ deve essere connesso, allora abbiamo dimostrato che esso è un albero. Cayley generalizzò il problema di individuare tutti i possibili idrocarburi saturi a quello di enumerare tutti i possibili alberi i cui vertici hanno grado 4 oppure 1. Egli trovò più semplice risolvere quest'ultimo problema e, inoltre, scoprì i grafi dei legami chimici di idrocarburi fino a quel momento ignoti che successivamente vennero scoperti.



Rappresentazione di alberi etichettati con colori per $n = 2, 3, 4$

Author: Julio Reis, immagine reperita sul sito https://en.wikipedia.org/wiki/File:Cayley's_formula_2-4.svg

Referenze

- M. Melo, lezioni di GE460 - Teoria dei Grafi, Università di Roma Tre, anno 2019/2020.
- J. A. Bondy – U. S. R. Murty, *Graph Theory*, Springer, 2008.

Grazie per l'attenzione